

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
15 February 2001 (15.02.2001)

PCT

(10) International Publication Number  
**WO 01/11041 A1**

(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C07K 14/495, A61K 38/18

(74) Agents: WEICKMANN, H. et al.; Kopernikusstrasse 9,  
D-81679 München (DE).

(21) International Application Number: PCT/EP00/07600

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 4 August 2000 (04.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
99115613.4 6 August 1999 (06.08.1999) EP

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): HY-GENE AG [CH/CH]; c/o Mäder + Baumgartner Treuhand AG, Schwanenfelsstrasse 10a, CH-8121 Neuhausen am Rheinfall (CH).

**Published:**

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): HÖTTEN, Gertrud [DE/DE]; Höllkeskampring 96, D-44625 Herne (DE). BECHTOLD, Rolf [DE/DE]; Carl-Zuckmayer-Strasse 21, D-69126 Heidelberg (DE). POHL, Jens [DE/DE]; Bastwaldstrasse 25, D-76707 Hambrücken (DE).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/11041 A1

(54) Title: MONOMERIC PROTEIN OF THE TGF- $\beta$  FAMILY

(57) Abstract: The present invention is concerned with proteins selected from the members of the TGF- $\beta$  superfamily, which are monomeric due to substitution or deletion of a cysteine which is responsible for dimer formation. The invention is also concerned with nucleic acids, encoding such monomeric proteins, vectors or host cells containing the nucleic acids as well as with pharmaceutical compositions comprising the proteins or nucleic acids encoding the proteins. The pharmaceutical compositions can be applied advantageously for all indications for which the respective dimeric proteins are useful.

## Monomeric Protein of the TGF- $\beta$ Family

### Description

5

The present invention concerns a biologically active protein from the TGF- $\beta$  superfamily, wherein this protein remains in monomeric form due to substitution or deletion of a cysteine which is responsible for the dimerization in the wild-type protein. Further the invention concerns a nucleic acid, which codes for a protein according to the invention, an expression vector containing such nucleic acid and a host cell, containing a corresponding nucleic acid or an expression vector, said nucleic acid being suitable for the expression of the protein. The invention also concerns a pharmaceutical composition containing the protein according to the invention or a nucleic acid coding therefor. The use of the pharmaceutical composition according to the invention concerns the prevention or treatment of all conditions which can also be treated with the dimeric form of the corresponding protein.

---

20 Many growth factors from the TGF- $\beta$  superfamily (Kingsley, Genes and Development 8, 133-146 (1994) as well as the references cited therein) are relevant for a wide range of medical treatment methods and applications which in particular concern promotion of cell proliferation and tissue formation, including wound healing and tissue reproduction. Such growth factors in particular comprise members of the TGF- $\beta$  (transforming growth factor, cf. e.g. Roberts and Sporn, Handbook of Experimental Pharmacology 95 (1990), page 419-472, editors: Sporn and Roberts), the DVR-group (Hötten et al., Biochem. Biophys. Res. Comm. 206 (1995), page 608-613 and further literature cited therein) including BMPs (bone morphogenetic protein, cf. e.g. Rosen and Thies, Growth Factors in Perinatal Development 30 (1993), page 39-58, editors: Tsang, Lemons and Balistreri) and GDFs (growth differentiation factors), the inhibin/activin (cf. e.g. Vale et al., The

Physiology of Reproduction, second edition (1994), page 1861-1878, editors: Knobil and Neill) and the GDNF protein family (Rosenthal, Neuron 22 (1999), page 201-203; Airaksinen et al. Mol Cell Neurosci 13 (1999), page 313-325). Although the members of the TGF- $\beta$  superfamily show high amino acid homologies in the mature part of the protein, in particular 5 conserved cysteines, they show considerable variations in their exact functions. Often individual growth factors of these families exhibit a plurality of functions at the same time, so that their application is of interest in various medical indications. Some of these multifunctional proteins also 10 have survival promoting effects on neurons in addition to functions such as e.g. regulation of the proliferation and differentiation in many cell types (Roberts and Sporn, supra; Sakurai et al., J. Biol. Chem. 269 (1994), page 14118-14122). Thus e.g. trophic effects on embryonic motoric and sensory neurons were demonstrated for TGF- $\beta$  in vitro (Martinou et al., Devl. Brain 15 Res. 52, page 175-181 (1990) and Chalazonitis et al., Dev. Biol. 152, page 121-132 (1992)). In addition, effects promoting survival are shown for dopaminergic neurons of the mid-brain for the proteins TGF- $\beta$ -1, -2, -3, activin A and GDNF (glial cell line-derived neurotrophic factor), a protein which has structural similarities to TGF- $\beta$  superfamily members, these 20 effects being not mediated via astrocytes (Krieglstein et al., EMBO J. 14, page 736-742 (1995)).

Interesting members of the TGF- $\beta$  superfamily or active variants thereof comprise the TGF- $\beta$  proteins like TGF- $\beta$ 1, TGF- $\beta$ 2, TGF- $\beta$ 3, TGF- $\beta$ 4, TGF- $\beta$ 5 25 (U.S. 5,284,763; EP 0376785; U.S. 4,886,747; DNA 7 (1988), page 1-8), EMBO J. 7 (1988), page 3737-3743), Mol. Endo. 2 (1988), page 1186-1195), J. Biol. Chem. 265 (1990), page 1089-1093), OP1, OP2 and OP3 proteins (U.S. 5,011,691, U.S. 5,652,337, WO 91/05802) as well as BMP2, BMP3, BMP4 (WO 88/00205, U.S. 5,013,649 and WO 89/10409, 30 Science 242 (1988), page 1528-1534), BMP5, BMP6 and BMP-7 (OP1) (Proc. Natl. Acad. Sci. 87 (1990), page 9841-9847, WO 90/11366), BMP8 (OP2) (WO 91/18098), BMP9 (WO 93/00432), BMP10 (WO 94/26893),

BMP11 (WO 94/26892), BMP12 (WO 95/16035), BMP13 (WO95/16035), BMP15 (WO 96/36710), BMP16 (WO 98/12322), BMP3b (Biochem. Biophys. Res. Comm. 219 (1996), page 656-662), GDF1 (WO 92/00382 and Proc. Natl. Acad. Sci. 88 (1991), page 4250-4254), GDF8 (WO 94/21681), GDF10 (WO95/10539), GDF11 (WO 96/01845), GDF5 (CDMP1, MP52) (WO 95/04819; WO96/01316; WO 94/15949, WO 96/14335 and WO 93/16099 and Nature 368 (1994), page 639-643), GDF6 (CDMP2, BMP13) (WO 95/01801, WO 96/14335 and WO95/16035), GDF7 (CDMP3, BMP12) (WO 95/01802 and WO 95/10635), GDF14 (WO 97/36926), GDF15 (WO 99/06445), GDF16 (WO 99/06556), 60A (Proc. Natl. Acad. Sci. 88 (1991), page 9214-9218), DPP (Nature 325 (1987), page 81-84), Vgr-1 (Proc. Natl. Acad. Sci. 86 (1989), page 4554-4558) Vg-1, (Cell 51 (1987), page 861-867), dorsalin (Cell 73 (1993), page 687-702), MIS (Cell 45 (1986), page 685-698), pCL13 (WO 97/00958), BIP (WO 94/01557), inhibin  $\alpha$ , activin  $\beta$ A and activin  $\beta$ B (EP 0222491), activin  $\beta$ C (MP121) (WO 96/01316), activin  $\beta$ E and GDF12 (WO 96/02559 and WO 98/22492), activin  $\beta$ D (Biochem. Biophys. Res. Comm. 210 (1995), page 581-588), GDNF (Science 260 (1993), page 1130-1132, WO 93/06116), Neurturin (Nature 384 (1996), page 467-470), Persephin (Neuron 20 (1998), page 245-253, WO 97/33911), Artemin (Neuron 21 (1998), page 1291-1302), Mic-1 (Proc. Natl. Acad. Sci USA 94 (1997), page 11514-11519), Univin (Dev. Biol. 166 (1994), page 149-158), ADMP (Development 121 (1995), page 4293-4301), Nodal (Nature 361 (1993), page 543-547), Screw (Genes Dev. 8 (1994), page 2588-2601). Other useful proteins include biologically active biosynthetic constructs including biosynthetic proteins designed using sequences from two or more known morphogenetic proteins. Examples of biosynthetic constructs are disclosed in U.S. 5,011,691 (e.g. COP-1, COP-3, COP-4, COP-5, COP-7 and COP-16). The disclosure of the cited publications including patents or patent applications are incorporated herein by reference.

The occurrence of proteins of the TGF- $\beta$  superfamily in various tissuous stages and development stages corresponds with differences with regard to their exact functions as well as target sites, life span, requirements for auxiliary factors, necessary cellular physiological environment and/or  
5 resistance to degradation.

The proteins of the TGF- $\beta$  superfamily exist as homodimers or heterodimers having a single disulfide bond. This disulfide bond is mediated by a specific and in most of the proteins conserved cysteine residue of the respective  
10 monomers. Up to now it was considered as indispensable for the biological activity that the protein is present in its dimeric form. Several publications indicated that biological activity can only be obtained for dimeric proteins and it was speculated that this dimer formation is important for further  
15 polymer formation of two or more dimers to achieve intercellular signal transmission by simultaneous binding to type I and type II receptors for the TGF- $\beta$  superfamily proteins on cells. It was assumed that only this simultaneous binding to both kinds of receptors would allow for effective intercellular signal transmission for the benefit of the patient (Bone, volume  
19 (1996), page 569-574).

20

A disadvantage of the use of these proteins as medicaments and their production is, that they are not readily obtainable in biologically active and sufficiently pure form by recombinant expression in prokaryots without intensive renaturation procedures.

25

Thus it was the object of the present invention to provide a simple and inexpensive possibility to reproducibly produce proteins exhibiting high biological activity, wherein this biological activity should essentially correspond to that of the dimers of the proteins of said families.

30

This object is solved according to the invention by a protein selected from the members of the TGF- $\beta$  protein superfamily, such protein being

necessarily monomeric due to substitution or deletion of a cysteine which is responsible for dimeric formation.

Surprisingly it has been found that the substitution or deletion of the cysteine, which normally effects the dimerization in the proteins, results upon expression and correct folding (proper formation of the intramolecular disulfide bridges) in a monomeric protein that retains the biological activity of the dimeric form. Even more surprisingly, it was found that at least some of the monomeric proteins show a higher activity, based on the weight of protein, than their respective dimeric forms. Apart from this improved biological activity an important advantage for the proteins according to the invention is that they can be expressed in a large amount in prokaryotic hosts and upon simple refolding of the monomers they are obtained in high purity and very high yield without the need to separate dimerized from non-dimerized (monomeric) protein. The findings of the present invention are very surprising since, as already mentioned above, it was common understanding that only a dimer of the morphogenetic proteins has biological activity. Despite this understanding the proteins according to the invention show an up to two-fold higher activity than that of the dimer on the basis of protein weight. The smaller size of the proteins of the invention, while maintaining the biological activity, can also be considered as advantageous, e.g. for applications concerning the brain since the monomeric protein can much easier pass the blood-brain-barrier than the dimeric form.

The proteins according to the invention encompass all proteins of the mentioned protein families that are normally present in dimeric form. Also parts of such proteins that retain substantial activity or fusion proteins or precursor forms of proteins shall be considered as encompassed by the present invention as well as biologically active naturally occurring or biosynthetic variants of TGF- $\beta$  superfamily proteins, as long as they show at least considerable biological activity.

In a preferred embodiment of the present invention the monomeric protein is a mature protein or a biologically active part or variant thereof. The term "biologically active part or variant thereof" is meant to define either fragments retaining activity, precursor proteins that are e.g. cleaved at the site of activity to the mature form or show biological activity themselves, or also variants that still maintain essentially the biological activity of the wild-type protein. Such variants preferably contain conservative amino acid substitutions, but especially at the N-terminal part of the mature proteins even considerable deletions or substitutions do not lead to a considerable loss of biological activity. It is well within the skill of the man in the art to determine whether a certain protein shows the required biological activity. Proteins showing at least 70% and preferably at least 80% homology to the mature wild-type proteins of the above referenced protein families should be understood as encompassed by the present invention, as long as they contain the deletion or substitution of a cysteine, as required for the proteins according the invention, and therefore do not form dimers.

It is especially preferred that proteins according to the invention contain at least the 7 cysteine region characteristic for the TGF- $\beta$  protein superfamily.

This specific 7 cysteine region is considered to be the most important part of the proteins in view of the biological activity. Therefore proteins retaining this critical region are preferred proteins according to the invention. It is disclosed in the state of the art which cysteine is responsible in a certain protein family or protein for dimer formation (see for example: Schlunegger & Grutter (1992) Nature 358, 430-434; Daopin et al., (1992) Science 257, 369-373 and Griffith et al., Proc. Natl. Acad. Sci. 93 (1996), page 878-883). This cysteine has to be deleted or substituted by another amino acid to form a protein according to the invention.

The 7 cysteine region is known for many proteins of the TGF- $\beta$  protein superfamily. In this region the respective location of the cysteine residues

to each other is important and is only allowed to vary slightly in order not to lose the biological activity. Consensus sequences for such proteins are known in the state of the art and all proteins complying with such consensus sequences are considered to be encompassed by the present invention.

In an especially preferred embodiment of the present invention the protein contains a consensus sequence according to the following sequence

C (Y)<sub>25-29</sub> C Y Y Y C (Y)<sub>25-35</sub> X C (Y)<sub>27-34</sub> C Y C (Formula I),

wherein C denotes cysteine, Y denotes any amino acid including cysteine and X denotes any amino acid except cysteine.

More preferably the protein according to the invention contains a consensus sequence according to the following sequence

C (Y)<sub>28</sub> C Y Y Y C (Y)<sub>30-32</sub> X C (Y)<sub>31</sub> C Y C (Formula II),  
wherein C, X and Y have the same meaning as defined above.

Even more preferably the protein according to the invention contains a consensus sequence according to the following sequence

C (X)<sub>28</sub> C X X X C (X)<sub>31-33</sub> C (X)<sub>31</sub> C X C (Formula III),  
wherein C and X have the same meaning as defined above.

In these consensus sequences especially preferred distances between the respective cysteine residues are contained, wherein also already the dimer forming cysteine is substituted by another amino acid. As with all proteins of said protein superfamily the location of and distance between the cysteines is more important than the identity of the other amino acids contained in this region. Therefore, the consensus sequence shows the respective location of the cysteines, but does not show the identity of the



other amino acids, since these other amino acids are widely variable in the proteins of the TGF- $\beta$  protein superfamily.

In a preferred embodiment of the present invention the monomeric protein  
5 according to the invention is a morphogenetic protein.

Most of the members of the TGF- $\beta$  protein superfamily are morphogenetic proteins that are useful for treatments where regulation of differentiation and proliferation of cells or progenitor cells is of interest. This can result in  
10 replacement of damaged and/or diseased tissue like for example skeletal (bone, cartilage) tissue, connective tissue, periodontal or dental tissue, neural tissue, tissue of the sensory system, liver, pancreas, cardiac, blood vessel and renal tissue, uterine or thyroid tissue etc. Morphogenetic proteins are often useful for the treatment of ulcerative or inflammatory  
15 tissue damage and wound healing of any kind such as enhanced healing of ulcers, burns, injuries or skin grafts. Especially preferred proteins according to the invention belong to the TGF- $\beta$ , BMP, GDF, activin or GDNF families. Several BMP proteins which were originally discovered by their ability to induce bone formation, have been described, as also indicated above.  
20 Meanwhile, several additional functions have been found as it is also true for members of the GDFs. These proteins show a very broad field of applications and especially are in addition to their bone and cartilage growth promoting activity (see for example: WO 88/00205, WO 90/11366, WO 91/05802) useful in periodontal disease, for inhibiting periodontal and tooth  
25 tissue loss, for sealing tooth cavities, for enhancing integration of a tooth in a tooth socket (see for example: WO 96/26737, WO 94/06399, WO 95/24210), for connective tissue such as tendon or ligament (see for example: WO 95/16035), for improving survival of neural cells, for inducing growth of neural cells and repairing neural defects, for damaged CNS tissue  
30 due to stroke or trauma (see for example: WO 97/34626, WO 94/03200, WO 95/05846), for maintaining or restoring sensory perception (see for example WO 98/20890, WO 98/20889), for renal failure (see for example:

WO 97/41880, WO 97/41881), for liver regeneration (see for example WO 94/06449), for regeneration of myocardium (see for example WO 98/27995), for treatment or preservation of tissues or cells for organ or tissue transplantation, for integrity of gastrointestinal lining (see for example  
5 WO 94/06420), for increasing progenitor cell population as for example hematopoietic progenitor cells by ex vivo stimulation (see for example WO 92/15323), etc. One preferred member of the GDF family is the protein MP52 which is also termed GDF-5 or CDMP-1. Applications for MP52 reflect several of the already described applications for the BMP/GDF family.  
10 MP52 is considered to be a very effective promoter of bone and cartilage formation as well as connective tissue formation (see for example WO 95/04819, Hötten et al., (1996), Growth Factors 13, 65-74, Storm et al., (1994) Nature 368, 639-643, Chang et al., (1994) J. Biol. Chem. 269 (45), 28227-28234). In this connection MP52 is useful for applications  
15 concerning the joints between skeletal elements (see for example Storm & Kingsley (1996) Development 122, 3969-3979). One example for connective tissue is tendon and ligament (Wolfman et al., (1997), J. Clin. Invest. 100, 321-330, Aspenberg & Forslund (1999), Acta Orthop Scand 70, 51-54, WO 95/16035). MP52 is also useful for tooth (dental and  
20 periodontal) applications (see for example WO 95/04819, WO 93/16099, Morotome et al. (1998), Biochem Biophys Res Comm 244, 85-90). MP52 is useful in wound repair of any kind. It is in addition very useful for promoting tissue growth in the neuronal system and survival of dopaminergic neurons, for example. MP52 in this connection is useful for  
25 applications in neurodegenerative diseases like e.g. Parkinson's disease and possibly also Alzheimer's disease for Huntington chorea tissues (see for example WO 97/03188, Kriegstein et al., (1995) J. Neurosci Res. 42, 724-732, Sullivan et al., (1997) Neurosci Lett 233, 73-76, Sullivan et al. (1998), Eur. J. Neurosci 10, 3681-3688). MP52 allows to maintain nervous  
30 function or to retain nervous function in already damaged tissues. MP52 is therefore considered to be a generally applicable neurotrophic factor. It is also useful for diseases of the eye, in particular retina cornea and optic

nerve (see for example WO 97/03188, You et al. (1999), Invest Ophthalmol Vis Sci 40, 296-311). The monomeric MP52 is expected to show all the already described activities of the dimeric form as well as some further described activities as described for the dimeric BMP/GDF family members.

5 It is expected to be for example also useful for increasing progenitor cell populations and for stimulating differentiation of progenitor cells ex vivo. Progenitor cells can be cells which take part in the cartilage formation process or hematopoietic progenitor cells. It is also useful for damaged or diseased tissue where a stimulation of angiogenesis is advantageous (see

10 for example: Yamashita et al. (1997), Exp Cell Res 235, 218-226).

An especially preferred protein according to the invention therefore is protein MP52 or a biologically active part or variant thereof. Like in the already above mentioned definition of these terms MP52 can e.g. be used

15 in its mature form, however, it can also be used as a fragment thereof at least containing the 7 cysteine region or also in a precursory form. Deviations at the N-terminal part of mature MP52 do not affect its activity to a considerable degree. Therefore, substitutions, deletions or additions on the N-terminal part of the proteins are still within the scope of the present

20 invention. It might be useful to add a peptide to the N-terminal part of the protein, e.g. for purification reasons. It might not be necessary to cleave off this added peptide after expression and purification of the protein. Additional peptides at the N- or C-terminal part of the protein may also serve for the targeting of the protein to special tissues such as nerve or

25 bone tissue or for the penetration of the blood/brain barrier. Generally, also fusion proteins of a monomeric protein according to the invention and another peptide or group are considered within the scope of the present invention, wherein these other peptides or groups are directing the localization of the fusion protein, e.g. because of an affinity to a certain

30 tissue type etc. Examples for such fusion proteins are described in WO 97/23612. The protein containing such addition will retain its biological

activity at least as long as such addition does not impair the formation of the biologically active conformation of the protein.

In an especially preferred embodiment of the present invention the proteins  
5 comprises the amino acid sequence according to SEQ.ID.NO.1 (DNA and  
protein sequence) and SEQ.ID.No.2 (protein sequence, only), respectively.  
SEQ.ID.NO.2 shows the complete protein sequence of the prepro protein  
of human MP52, as already disclosed in WO 95/04819. The start of the  
mature protein lies preferably in the area of amino acids 352-400, especially  
10 preferred at amino acids 381 or 382. Therefore, the mature protein  
comprises amino acids 381-501 or 382-501. The first alanine of the mature  
protein can be deleted and the mature protein then preferably comprises  
amino acids 383-501. The cysteine at position 465 that is present in the  
already described dimeric MP52 protein is according to the invention either  
15 deleted or substituted by another amino acid. This deletion or substitution  
is represented by Xaa at the respective position in SEQ.ID.Nos.1 and 2.

The activin/inhibin family proteins are of interest for applications related to  
contraception, fertility and pregnancy (see for example WO 94/19455, U.S.  
20 5,102,868). They are also of interest for applications like repair or  
prevention of diseases of the nervous system, they can be used in the  
repair of organ tissue such as liver and even in bone and cartilage, too. In  
this connection MP121 (activin  $\beta$ C) is especially useful in applications for  
growth or regeneration of damaged and/or diseased tissue, especially the  
25 liver tissue, neural tissue, skeletal tissue (see for example WO 96/01316,  
WO 98/22492 and WO 97/03188). MP121 is known to be predominantly  
expressed in the liver whereby the mRNA is markedly reduced after partial  
hepatectomy. MP121 is expected to regulate the liver mass (Zhang et al.,  
Endocrine Journal 44 (1997), page 759-764). The monomeric MP121  
30 shows all the already described activities of the dimeric form as well as  
some further described activities as described for the dimeric TGF- $\beta$   
superfamily members. It is for example also expected to be useful in

treatment of ulceration (for example stomach ulceration) and useful for integrity of gastrointestinal lining and for stimulating differentiation of progenitor cells ex vivo, treatment or preservation of mammalian tissue or cells, e.g. for organ or tissue transplantation.

5

A further preferred protein according to the invention therefore is MP121, a member of the activin/inhibin protein family. Also for this protein a biologically active part or variant thereof is encompassed by the present invention according to the above defined rules. An especially preferred  
10 embodiment is shown in SEQ.ID.NO.3 (DNA and protein sequence) and SEQ.ID.NO.4 (protein sequence, only) respectively. SEQ.ID.NO.4 shows the complete amino acid sequence of the prepro protein of human MP121, that has already been disclosed in WO 96/01316. The start of the mature protein lies preferably between amino acids 217 and 247, most preferred  
15 at amino acid 237. A preferred mature protein therefore comprises the mature part of the protein starting at amino acid 237 and ending at amino acid 352. However, also the precursor protein comprising the whole shown amino acid sequence is encompassed by the present invention. The cysteine at position 316 is according to the invention either deleted or substituted  
20 by another amino acid, being represented by Xaa in SEQ.ID.Nos.3 and 4.

The amino acid by which the cysteine residue effecting the dimerization is substituted can be selected by any amino acid that does not impair the formation of a biologically active conformation. The amino acid is preferably  
25 selected from the group of alanine, serine, threonine, leucine, isoleucine, glycine and valine.

The proteins according to the invention are in summary characterized by the absence of the cysteine residue in the amino acid sequence responsible for  
30 the dimer formation. This absence can be effected by substitution of this cysteine by another amino acid or by deletion. In case of deletion, however, it must be assured for the protein that the formation of the biologically

active conformation is not hindered. The same is true for the selection of the substitution amino acid, wherein it is preferred to use an amino acid which has a form similar to cysteine.

- 5 The monomeric proteins according to the invention can be easily produced, in particular by expression in prokaryots and renaturation according to known methods. It is advantageous that the protein can be obtained in exceedingly biologically active form. The proteins exhibit in monomeric form about the same activity as the dimer so that based on the amount of active  
10 substance only the same amount of the monomeric protein has to be used in order to obtain the same positive biological effects.

A further subject matter of the present invention is a nucleic acid encoding a protein according to the invention. It is obvious that the nucleic acid has  
15 to have such a sequence that a deletion or substitution of the cysteine responsible for the dimer formation is achieved. The nucleic acid can be a naturally occurring nucleic acid, but also a recombinantly produced or processed nucleic acid. The nucleic acid can be both a DNA sequence and an RNA sequence, as long as the protein according to the invention can be  
20 obtained from this nucleic acid upon expression in a suitable system.

In a preferred embodiment of the invention the nucleic acid is a DNA sequence. This DNA sequence in an especially preferred embodiment of the invention comprises a sequence as shown in SEQ.ID.NO.1 and  
25 SEQ.ID.NO.3, respectively, or parts thereof. SEQ.ID.NO.1 shows a nucleic acid encoding MP52, wherein the codon for the cysteine responsible for the dimer formation is replaced by another codon which does not encode cysteine or deleted. This substitution or deletion is shown as "nnn" in the sequence protocols. SEQ.ID.NO.3 shows a nucleic acid encoding MP121,  
30 wherein also the codon for the cysteine effecting the dimer formation is replaced by a respective different codon or deleted. Instead of the complete

sequences of SEQ.ID.NOs.1 or 3 also parts can be used that encode the mature proteins or fragments also described above.

It is preferred in the framework of the present invention that the nucleic acid apart from the coding sequences also contains expression control sequences. Such expression control sequences are known to the man skilled in the art and serve to control the expression of the encoded protein in a host cell. The host cell does not have to be an isolated cell, moreover, the nucleic acid can be expressed in the patient in vivo in the target tissue. This can be done by inserting the nucleic acid into the cell genome, however, it is also possible to transform host cells with expression vectors containing a nucleic acid according to the invention. Such expression vectors are a further subject matter of the present invention, wherein the nucleic acid is inserted in a suitable vector system, the vector system being selected according to the desired expression of the protein. The vector system can be a eukaryotic vector system, but - in the framework of the present invention - it is preferably a prokaryotic vector system, with which the proteins can be produced in prokaryotic host cells in a particularly easy and pure manner. In addition, the expression vector can be a viral vector.

Also host cells in turn are a further subject matter of the present invention. The host cells are characterized in that they contain a nucleic acid according to the invention or an expression vector according to the invention and that they are able to use the information present in the nucleic acids and in the expression vector, respectively, for the expression of a monomeric protein according to the invention.

Although in the framework of the present invention also eukaryotic host cells are suitable for the production of the protein, it is, as mentioned already several times above, particularly advantageous that the protein according to the invention can be produced in prokaryotic host cells, which therefore represent a preferred embodiment of the present invention.

After such preferred expression in prokaryotic host cells the protein is purified and renatured according to known methods, thereby effecting intramolecular cystine bridge formation.

5 Since, however, not only in vitro production of the monomeric protein is possible, but also in vivo expression of a nucleic acid according to the invention, a further preferred embodiment is a eukaryotic host cell, and especially a eukaryotic host cell containing the DNA in its genome, or as an expression vector. Such host cell can also be useful for application to an  
10 individual in need of morphogenic treatment.

Further subject matters of the present application are pharmaceutical compositions comprising at least one monomeric protein according to the invention or at least one nucleic acid encoding for such a protein or at least  
15 one corresponding expression vector, or at least one eukaryotic host cell expressing the monomeric protein.

The protein itself, but also a nucleic acid according to the invention, an expression vector or a host cell can be considered to be advantageous as  
20 active substances in a pharmaceutical composition. Also combinations of monomeric proteins, with either biological activities in the same or different applications, can be used in preferred pharmaceutical compositions. Especially preferred for neuronal applications are combinations of MP52 with other TGF- $\beta$  superfamily proteins, both in monomeric form, like for  
25 example with GDNF (see WO 97/03188). Also preferred for neuronal applications are combinations of TGF- $\beta$  with GDNF, both in monomeric form. Also for applications concerning cartilage and/or bone the combination of several monomeric proteins might be useful, like MP52 with a protein of TGF- $\beta$  (see e.g. WO 92/09697) or MP52 with a cartilage  
30 maintenance-inducing protein such as BMP-9 (see e.g. WO 96/39170). When a nucleic acid or an expression vector is used, however, it has to be ensured that when administering to the patient there has to be an



environment in which the nucleic acid and the expression vector, respectively, can be expressed and the protein according to the invention can be produced in vivo at the site of action. The same applies accordingly to the host cell according to the invention. When using expression vectors  
5 or host cells it is also possible that they encode more than one monomeric protein of the invention to produce a combination of two or more monomeric proteins.

It is advantageous to both the protein and the nucleic acid or the expression  
10 vector or the host cell when they are applied in and/or on a biocompatible matrix. The matrix material can be transplanted into the patient, e.g. surgically, wherein the protein either is effective on the surface of the matrix material or the protein or the DNA encoding the protein can be slowly released from the matrix material and then be effective over a long  
15 period of time. Additionally it is possible and advantageous to use a biodegradable matrix material in the pharmaceutical composition, wherein this material preferably dissolves during the protein induced tissue formation so that a protein or a nucleic acid contained therein is released and the newly formed tissue replaces the matrix material.

20 Finally, in case of applications relating to bone formation, it is advantageous to use a matrix material which is itself e.g. osteogenically active. By using such a matrix material it becomes possible to achieve a synergistic effect of protein and matrix material and to effect a particularly rapid and effective  
25 bone formation.

An especially preferred matrix material that can be used according to the invention is a matrix material as described in U.S. 5,231,169 and U.S. 5,776,193 and especially for applications like spinal fusion.

30 When using a combination of a matrix material and protein and/or nucleic acid and/or expression vector, it is preferable to sterilize such a combination

prior to its use. The matrix and the morphogenetic protein can be separately sterilized and then combined, but it is preferred to terminally sterilize the device consisting of matrix and morphogenetic protein. Terminal sterilization can be achieved by ionizing radiation as already described for dimeric proteins (U.S. 5,674,292) but it is also advantageous to use ethylene oxide.

Of course this invention also comprises pharmaceutical compositions containing further substances like e.g. pharmacologically acceptable auxiliary and carrier substances. However, the protein according to the invention, also in case a matrix material is used, does not necessarily have to be used together with this matrix material, but can also be administered systemically, wherein it concentrates preferably in the surrounding of an implanted matrix material.

For some applications the protein according to the invention and the nucleic acid forming this protein, respectively or the expression vector or host cell can preferably be present in an injectable composition. Implants are not necessary or possible for every form of application of the proteins according to the invention. However, it is also possible to provide an implantable vessel or an implantable micropump containing for example semipermeable membranes in which the protein according to the invention or the nucleic acid generating it is contained, from which either one is slowly released over a prolonged period of time. The pharmaceutical composition according to the invention can also contain other vehicles which make it possible that the proteins or the nucleic acids or the expression vectors encoding these proteins be transported to the site of activity and released there, wherein e.g. liposomes or nanospheres can be used. In principle, it is also possible to apply host cells, like e.g. implanted embryonic cells expressing the proteins. Cells transfected with recombinant DNA may be encapsulated prior to implantation. Any other practicable but herein not explicitly described form of application of the pharmaceutical composition according to the invention and their corresponding manufacture are also comprised by the

present invention, as long as they contain a protein according to the invention or a nucleic acid or an expression vector coding therefor, or a host cell expressing it.

5 Although the indications shall not be restricted herein and all indications exhibiting the dimeric form of the protein according to the invention are also comprised, in the following types of application for the compositions according to the invention are listed which are considered to be particularly preferred indications for the proteins of the present invention. On the one  
10 hand, there is the prevention or therapy of diseases associated with bone and/or cartilage damage or affecting bone and/or cartilage disease, or generally situations, in which cartilage and/or bone formation is desirable or for spinal fusion, and on the other hand, there is prevention or therapy of damaged or diseased tissue associated with connective tissue including  
15 tendon and/or ligament, periodontal or dental tissue including dental implants, neural tissue including CNS tissue and neuropathological situations, tissue of the sensory system, liver, pancreas, cardiac, blood vessel, renal, uterine and thyroid tissue, skin, mucous membranes, endothelium, epithelium, for promotion or induction of nerve growth, tissue  
20 regeneration, angiogenesis, wound healing including ulcers, burns, injuries or skin grafts, induction of proliferation of progenitor cells or bone marrow cells, for maintenance of a state of proliferation or differentiation for treatment or preservation of tissue or cells for organ or tissue transplantation, for integrity of gastrointestinal lining, for treatment of  
25 disturbances in fertility, contraception or pregnancy.

Diseases concerning sensory organs like the eye are also to be included in the preferred indication of the pharmaceutical composition according to the invention. As neuronal diseases again Parkinson's and Alzheimer's diseases  
30 can be mentioned as examples.

The pharmaceutical compositions according to the invention can be used in any desired way, the pharmaceutical compositions are formulated preferably for surgical local application, topical or systemic application. Auxiliary substances for the individual application form can of course be present in the pharmaceutical composition according to the invention. For some applications it can be advantageous to add some further substances to the pharmaceutical composition as for example Vitamin D (WO 92/21365), parathyroid hormone related peptide (WO 97/35607), chordin (WO 98/21335), anti-fibrinolytic agent (EP 535091), anti-metabolites (WO 95/09004), alkyl cellulose (WO 93/00050), mannitol (WO 98/33514), sugar, glycine, glutamic acid hydrochloride (U.S. 5,385,887), antibiotics, antiseptics, amino acids and/or additives which improve the solubility or stability of the monomeric morphogenetic protein as for example nonionic detergents (e.g. Tween 80), basic amino acids, carrier proteins (e.g. serum albumin), full length propeptides of the TGF- $\beta$  superfamily or parts thereof.

As can be already gathered from the description of proteins, nucleic acids and pharmaceutical compositions, the proteins according to the invention and respective nucleic acids, which provide for an expression of the proteins at the site of activity, can advantageously be applied in all areas for which also the dimeric forms of the proteins, as described, can be applied. In the framework of the present invention therefore a further subject matter is the use of a pharmaceutical composition according to the present invention for the treatment or prevention of any indications of the dimeric forms of the proteins according to the invention.

Herein it is again possible to conduct surgical operations and to implant the pharmaceutical composition (in particular contained on a matrix material), an administration in liquid or otherwise suitable form via, e.g. injection or oral administration seems to be as suitable as a topical application for e.g. tissue regeneration.

Fig. 1A shows a two dimensional graph of the conformation of recombinantly produced dimeric MP52 with the deleted first alanine. In this figure the 7 cysteine bridges contained in a dimer are shown, wherein there are 3 intramolecular cystine bridges per monomer unit and 1 intermolecular cystine bridge connecting both monomers. Fig. 1B shows the monomeric protein according to the invention wherein the cysteine of the amino acid sequence of MP52 has been replaced by X that denotes any amino acid except cysteine.

## 10 Examples

The examples shall explain the invention more illustratively and are not considered to be restrictive.

### 15 Expression of monomeric MP52 in E. coli, refolding , purification and possible test systems to determine the biological activity

The mature part of human MP52 was mutated by converting Cysteine to Alanine at position 465 of SEQ ID NO 2. For this purpose the nucleotides at the position 2032 and 2033 of SEQ ID NO 1 were converted from TG to GC. The substitution of the nucleotides was performed by using the prokaryotic plasmid pBP2MP52m and a PCR method.

The vector pBP2MP52m is a derivative of the pBR322 plasmid containing an ampicillin resistance gene, a T7-promoter followed by a ribosomal binding site, a start codon, the mature part of MP52 (coding for the amino acids 382-501 in SEQ ID NO. 2 with a Cysteine at position 465), stop codons in each reading frame and a terminator. The plasmid was deposited at the DSM (DSM 10029, 2nd June 1995).

The mutation was performed by using the QuickChange™ Site-Directed Mutagenesis Kit with the PfuTurbo™ DNA polymerase and the Dpn I endonuclease from Stratagene (Catalog #200518) according to the

- 21 -

- instruction manual of the manufacturer. The oligonucleotide CCACACCACCCACCGCCTGTGTGCCCACGC (SEQ ID NO. 5) and the oligonucleotide GCGTGGGCACACAGGCGGTGGGTGGTGTGG (SEQ ID NO. 6) were purified by polyacrylamide gel electrophoresis and used as the mutagenic primers. The resulting plasmid pBP2MP52<sub>ALA</sub> contained after the start codon the nucleotide sequence coding for the amino acids 382-501 in SEQ ID NO. 2 with an Alanine at position 465 which was verified by sequencing.
- 10 The expression of monomeric MP52 can be induced by providing a source of T7 RNA polymerase. Using the bacterial strain BL21(DE3)pLysS (Novagen) transformed with the plasmid pBP2MP52<sub>ALA</sub> and inducing the T7 RNA polymerase gene according to the manufacturers instructions with IPTG, monomeric MP52 can be expressed in inclusion bodies which can be
- 15 isolated according to standard procedures. Further purification was done on a reversed phase column (Nucleosil 300-7C4, Machery-Nagel) with a 0 to 50% buffer B gradient (buffer A: 0.1% TFA in water, buffer B: 90% acetonitrile, 0.1 % TFA) in 50 minutes (flow rate: 2 ml/min). The fractions containing monomeric MP52 were pooled, lyophilized and stored at -70°C.
- 20 MP52 was solubilized in a denaturing buffer (6M Urea, 500 mM NaCl, 10 mM DTT, 1 mM EDTA , 20 mM Tris pH 8.3) and added for refolding in 9 times of a 50 mM Glycine-buffer system (pH 9.8) containing CHAPS (20 mM), NaCl (500 mM) and 3 mM GSSG and gently stirred at 4°C. After approximately 24 hours the sample was diluted 2.8 times with 14 mM
- 25 NaH<sub>2</sub>PO<sub>4</sub> and subjected to isoelectric precipitation. After centrifugation the precipitate was dissolved in 0.1 % TFA and the folded monomer was further purified by reversed phase HPLC. For this purpose MP52 was loaded on a column (Aquapore Octyl 20 micron, Applied Biosystems) equilibrated with 25% buffer B (buffer A: 0.1% TFA in water, buffer B: 90% acetonitrile, 0.1 % TFA). Monomeric MP52 was eluted with a 25-60%
- 30 buffer B gradient in 70 minutes (flow rate 3 ml/min). Fractions containing purified refolded monomeric MP52 were pooled, lyophilized, stored at -70°C

and can be used in biological activity studies. Useful for biological activity studies of monomeric MP52 in the field of cartilage and bone induction are for example the measurement of increasing ALP activity (Takuwa et al., Am. J. Physiol. 257, E797-E803, 1989) using osteoprogenitor-like ROB-C26  
5 cells (Yamaguchi et al., Calcif. Tissue Int. 49, 221-225, 1991) as described in WO95/04819 or ATDC5 cells (Riken Gene Bank, RCB 0565, embryonic cells which differentiate like cartilage cells). Useful activity tests showing the neurological capacities of refolded monomeric MP52 are the measurement of increased survival of dopaminergic neurons as described  
10 for example by Kriegstein et al. (J. Neuroscience Res. 42, 724-732, 1995) or Sullivan et al. (Neuroscience Letters 233, 73-76, 1997) or the outgrowth of nerve fibers from embryonic retina are described for example in the WO 97/03188. The angiogenic potential of folded monomeric MP52 can be verified for example in an *in vivo* corneal micropocket model. Briefly,  
15 Hydron can be used as a slow release pellet adapted from D'Amato et al (Proc. Natl. Acad. Sci. USA 91 (1994) 4082-4085). About 2  $\mu$ g of monomeric MP52 can be dissolved in 20  $\mu$ l of 50% acetonitrile, stabilized with 5 mg sucralfate (sucrose aluminum sulfate; Bukh Meditec, Copenhagen) and mixed with 20  $\mu$ l of 12% (wt/vol) Hydron NCC (Interferon  
20 Science, New Brunswick, NJ) in ethanol. The mixture can be pipetted onto Teflon pegs and dried to produce a pellet. Subsequently the pellet can be implanted into corneal micropockets of the eye of a male New Zealand White rabbit. For this purpose an intrastromal tunnel, starting from a perpendicular linear incision in the pupillary plane and ending 2.75 mm prior  
25 to the limbus can be created. As a postsurgical care one single dose of erythromycin ointment should be given onto the surface of the cornea. About 10 days after implantation new capillaries which grow toward the implantation site should be seen. The outgrowth towards the pellet can reflect the chemotactic response of the endothelial cells to the released  
30 MP52 and can be quantified by measuring the vascularized area and the number of vessels at the limbus .

Monomeric MP52 isolated and purified from inclusion bodies but not subjected to the refolding process can be used as a negative control. Dimeric MP52 can be used as the positive control.

5    **Expression of monomeric MP121 in E.coli, refolding, purification and possible test systems to determine the biological activity**

The mature part of human MP121 was mutated by converting Cysteine to Alanine at position 316 of SEQ ID NO 4. For this purpose the nucleotides  
10    at position 1073 and 1074 of SEQ ID NO. 3 were converted from TG to GC. The substitution of the nucleotides was performed essentially as described for MP52 in the above example with differences in the plasmid and the mutagenic primers. The pBP4MP121m plasmid was used which differs from the pBP2MP52m plasmid by containing the mature part of  
15    MP121 (coding for the amino acids 237-352 in SEQ ID NO. 4 with a Cysteine at position 316) instead of the mature part of MP52 and by harboring a tetracycline resistance gene instead of the ampicillin resistance gene. As mutagenic primers the oligonucleotide CTGGAGGGGGCTCAGCCTGTGTACCCACGG (SEQ ID NO. 7) and the  
20    oligonucleotide CCGTGGGTACACAGGCTGAGCCCCCTCCAG (SEQ ID NO. 8) were used.

The resulting plasmid pBP4MP121<sub>ALA</sub> contained after the start codon the nucleotide sequence coding for the amino acids 237-352 in SEQ ID NO. 4  
25    with an Alanine at position 316 which was verified by sequencing.

The expression of monomeric MP121, isolation of inclusion bodies and purification was performed essentially as described for MP52 in the above example, with the exception that another E.coli strain, the HMS 174 (DE3)  
30    (Novagen), was used.



For biological activity studies of refolded monomeric MP121 showing the neurological capacities the same assays as mentioned for MP52 can be used. The influence of refolded monomeric MP121 on liver growth can be investigated for example on rat hepatocytes in primary cultures by inhibiting  
5 EGF induced DNA synthesis. Briefly, hepatocytes can be isolated from rat (Wistar) liver and cultured according to Yasuda et al. (J. Clin. Invest. Vol. 92, 1491-1496 (1993)). The cells have to be washed prior to incubation with fresh serum-free medium containing 0.1 nM insulin, 0.1% BSA and 1 nM EGF. The solubilized refolded monomeric MP121 can be added to the  
10 medium at various concentrations and subsequently the hepatocytes have to be incubated for 72h including 0.5  $\mu$ Ci [ $^3$ H]Thymidine/ml for the last 24 hours as described by Mead & Fausto (Proc. Natl. Acad. Sci. USA 86, 1558-15562 (1989)). [ $^3$ H]Thymidine incorporation into trichloroacetic acid-precipitable material can be measured as described by McNiel et al. (J. Cell  
15 Biol. 101, 372-379 (1985)).

Stimulation of hematopoiesis can be determined for examples on bone marrow aspirate cultures as described in example 8 of WO98/22492.

## SEQUENCE LISTING

5     <110> HyGene AG

          <120> Monomeric Protein of the TGF-beta Family

          <130> 20780pwomd

10           <140>

          <141>

          <150> EP99115613

15           <151> 1999-08-06

          <160> 8

          <170> PatentIn Ver. 2.1

20           <210> 1

          <211> 2703

          <212> DNA

          <213> Homo sapiens

25           <220>

---

          <221> CDS

          <222> (640) .. (2142)

30     <400> 1

ccatggcctc gaaagggcag cggtgatttt ttccacataa atatatcgca cttaaatgag 60

tttagacagc atgacatcag agagtaatta aattggtttg gggttgaatt ccgtttccaa 120

35     ttcctgagtt caggtttgta aaagattttt ctgagcacct gcaggcctgt gagtgtgtgt 180

gtgtgtgtgt gtgtgtgtgt gtgtgtgtga agtattttca ctggaaagga ttcaaaacta 240

40     gggggaaaaa aaaactggag cacacaggca gcattacgcc attcttcctt cttggaaaaa 300

- 26 -

tccctcagcc ttatacaagc ctccttcaag ccctcagtc gttgtgcagg agaaagggg 360  
 cggttggett tctcctttca agaacgagtt attttcagct gctgactgga gacggtgcac 420  
 5 gtctggatac gagagcattt ccactatggg actggataca aacacacacc cggcagactt 480  
 caagagtctc agactgagga gaaagccttt ccttctgctg ctactgctgc tgccgctgct 540  
 tttgaaagtc cactcctttc atgggttttc ctgccaaacc agaggcacct ttgctgctgc 600  
 15 cgctgttctc tttggtgtca ttcagcggtt ggccagagg atg aga ctc ccc aaa 654  
 Met Arg Leu Pro Lys  
 1 5  
 20 ctc ctc act ttc ttg ctt tgg tac ctg gct tgg ctg gac ctg gaa ttc 702  
 Leu Leu Thr Phe Leu Leu Trp Tyr Leu Ala Trp Leu Asp Leu Glu Phe  
 10 15 20  
 atc tgc act gtg ttg ggt gcc cct gac ttg ggc cag aga ccc cag ggg 750  
 25 Ile Cys Thr Val Leu Gly Ala Pro Asp Leu Gly Gln Arg Pro Gln Gly  
 25 30 35  
 acc agg cca gga ttg gcc aaa gca gag gcc aag gag agg ccc ccc ctg 798  
 Thr Arg Pro Gly Leu Ala Lys Ala Glu Ala Lys Glu Arg Pro Pro Leu  
 30 40 45 50  
 gcc cgg aac gtc ttc agg cca ggg ggt cac agc tat ggt ggg ggg gcc 846  
 Ala Arg Asn Val Phe Arg Pro Gly Gly His Ser Tyr Gly Gly Gly Ala  
 55 60 65  
 35 acc aat gcc aat gcc agg gca aag gga ggc acc ggg cag aca gga ggc 894  
 Thr Asn Ala Asn Ala Arg Ala Lys Gly Gly Thr Gly Gln Thr Gly Gly  
 70 75 80 85  
 40 ctg aca cag ccc aag aag gat gaa ccc aaa aag ctg ccc ccc aga ccg 942  
 Leu Thr Gln Pro Lys Lys Asp Glu Pro Lys Lys Leu Pro Pro Arg Pro  
 90 95 100

- 27 -

ggc ggc cct gaa ccc aag cca gga cac cct ccc caa aca agg cag gct 990  
 Gly Gly Pro Glu Pro Lys Pro Gly His Pro Pro Gln Thr Arg Gln Ala  
 105 110 115

5 aca gcc cgg act gtg acc cca aaa gga cag ctt ccc gga ggc aag gca 1038  
 Thr Ala Arg Thr Val Thr Pro Lys Gly Gln Leu Pro Gly Gly Lys Ala  
 120 125 130

10 ccc cca aaa gca gga tct gtc ccc agc tcc ttc ctg ctg aag aag gcc 1086  
 Pro Pro Lys Ala Gly Ser Val Pro Ser Ser Phe Leu Leu Lys Lys Ala  
 135 140 145

15 agg gag ccc ggg ccc cca cga gag ccc aag gag cgg ttt cgc cca ccc 1134  
 Arg Glu Pro Gly Pro Pro Arg Glu Pro Lys Glu Pro Phe Arg Pro Pro  
 150 155 160 165

ccc atc aca ccc cac gag tac atg ctc tcg ctg tac agg acg ctg tcc 1182  
 Pro Ile Thr Pro His Glu Tyr Met Leu Ser Leu Tyr Arg Thr Leu Ser  
 170 175 180

20 gat gct gac aga aag gga ggc aac agc agc gtg aag ttg gag gct ggc 1230  
 Asp Ala Asp Arg Lys Gly Gly Asn Ser Ser Val Lys Leu Glu Ala Gly  
 185 190 195

25 ctg gcc aac acc atc acc agc ttt att gac aaa ggg caa gat gac cga 1278  
 Leu Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys Gly Gln Asp Asp Arg  
 200 205 210

---

30 ggt ccc gtg gtc agg aag cag agg tac gtg ttt gac att agt gcc ctg 1326  
 Gly Pro Val Val Arg Lys Gln Arg Tyr Val Phe Asp Ile Ser Ala Leu  
 215 220 225

35 gag aag gat ggg ctg ctg ggg gcc gag ctg cgg atc ttg cgg aag aag 1374  
 Glu Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg Ile Leu Arg Lys Lys  
 230 235 240 245

ccc tcg gac acg gcc aag cca gcg gcc ccc gga ggc ggg cgg gct gcc 1422  
 Pro Ser Asp Thr Ala Lys Pro Ala Ala Pro Gly Gly Gly Arg Ala Ala  
 250 255 260

40 cag ctg aag ctg tcc agc tgc ccc agc ggc cgg cag cgg gcc tcc ttg 1470  
 Gln Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg Gln Pro Ala Ser Leu  
 265 270 275

- 28 -

ctg gat gtg cgc tcc gtg cca ggc ctg gac gga tct ggc tgg gag gtg 1518  
 Leu Asp Val Arg Ser Val Pro Gly Leu Asp Gly Ser Gly Trp Glu Val  
 280 285 290

5 ttc gac atc tgg aag ctc ttc cga aac ttt aag aac tcg gcc cag ctg 1566  
 Phe Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys Asn Ser Ala Gln Leu  
 295 300 305

10 tgc ctg gag ctg gag gcc tgg gaa cgg ggc agg gcc gtg gac ctc cgt 1614  
 Cys Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg Ala Val Asp Leu Arg  
 310 315 320 325

15 ggc ctg ggc ttc gac cgc gcc gcc cgg cag gtc cac gag aag gcc ctg 1662  
 Gly Leu Gly Phe Asp Arg Ala Ala Arg Gln Val His Glu Lys Ala Leu  
 330 335 340

ttc ctg gtg ttt ggc cgc acc aag aaa cgg gac ctg ttc ttt aat gag 1710  
 Phe Leu Val Phe Gly Arg Thr Lys Lys Arg Asp Leu Phe Phe Asn Glu  
 345 350 355

20 att aag gcc cgc tct ggc cag gac gat aag acc gtg tat gag tac ctg 1758  
 Ile Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr Val Tyr Glu Tyr Leu  
 360 365 370

25 ttc agc cag cgg cga aaa cgg cgg gcc cca ctg gcc act cgc cag ggc 1806  
 Phe Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu Ala Thr Arg Gln Gly  
 375 380 385

30 aag cga ccc agc aag aac ctt aag gct cgc tgc agt cgg aag gca ctg 1854  
 Lys Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys Ser Arg Lys Ala Leu  
 390 395 400 405

35 cat gtc aac ttc aag gac atg ggc tgg gac gac tgg atc atc gca ccc 1902  
 His Val Asn Phe Lys Asp Met Gly Trp Asp Asp Trp Ile Ile Ala Pro  
 410 415 420

ctt gag tac gag gct ttc cac tgc gag ggg ctg tgc gag ttc cca ttg 1950  
 Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu  
 425 430 435

40 cgc tcc cac ctg gag ccc acg aat cat gca gtc atc cag acc ctg atg 1998  
 Arg Ser His Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met  
 440 445 450

- 29 -

aac tcc atg gac ccc gag tcc aca cca ccc acc nnn tgt gtg ccc acg 2046  
 Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr Xaa Cys Val Pro Thr  
 455 460 465

5 . cgg ctg agt ccc atc agc atc ctc ttc att gac tct gcc aac aac gtg 2094  
 Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val  
 470 475 480 485

10 gtg tat aag cag tat gag gac atg gtc gtg gag tcg tgt ggc tgc agg 2142  
 Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys Arg  
 490 495 500

tagcagcact ggccctctgt cttcctgggt ggcacatccc aagagccctt tcctgcactc 2202

15 ctggaatcac agaggggtca ggaagctgtg gcaggagcat ctacacagct tgggtgaaag 2262  
 gggattccaa taagcttgct cgctctctga gtgtgacttg ggctaaaggc ccccttttat 2322  
 ccacaagttc cctggctga ggattgctgc ccgtctgctg atgtgaccag tggcaggcac 2382  
 20 aggtccaggg agacagactc tgaatgggac tgagtcccag gaaacagtgc tttccgatga 2442  
 gactcagccc accatttctc ctcacctggg cttctcagc ctctggactc tcctaagcac 2502

25 ctctcaggag agccacaggt gccactgcct cctcaaatca catttggtgcc tggtgacttc 2562  
 ctgtccctgg gacagttgag aagctgactg ggcaagagtg ggagagaaga ggagagggct 2622

---

30 tggatagagt tgaggagtgt gaggtgtta gactgttaga tttaaata tttgatgag 2682  
 ataaaaagca aaactgtgcc t 2703

35 <210> 2  
 <211> 501  
 <212> PRT  
 <213> Homo sapiens

40 <400> 2  
 Met Arg Leu Pro Lys Leu Leu Thr Phe Leu Leu Trp Tyr Leu Ala Trp  
 1 5 10 15  
 Leu Asp Leu Glu Phe Ile Cys Thr Val Leu Gly Ala Pro Asp Leu Gly  
 20 25 30  
 45

- 30 -

Gln Arg Pro Gln Gly Thr Arg Pro Gly Leu Ala Lys Ala Glu Ala Lys  
 35 40 45

Glu Arg Pro Pro Leu Ala Arg Asn Val Phe Arg Pro Gly Gly His Ser  
 5 50 55 60

Tyr Gly Gly Gly Ala Thr Asn Ala Asn Ala Arg Ala Lys Gly Gly Thr  
 65 70 75 80

Gly Gln Thr Gly Gly Leu Thr Gln Pro Lys Lys Asp Glu Pro Lys Lys  
 10 85 90 95

Leu Pro Pro Arg Pro Gly Gly Pro Glu Pro Lys Pro Gly His Pro Pro  
 15 100 105 110

Gln Thr Arg Gln Ala Thr Ala Arg Thr Val Thr Pro Lys Gly Gln Leu  
 115 120 125

Pro Gly Gly Lys Ala Pro Pro Lys Ala Gly Ser Val Pro Ser Ser Phe  
 20 130 135 140

Leu Leu Lys Lys Ala Arg Glu Pro Gly Pro Pro Arg Glu Pro Lys Glu  
 145 150 155 160

Pro Phe Arg Pro Pro Pro Ile Thr Pro His Glu Tyr Met Leu Ser Leu  
 25 165 170 175

Tyr Arg Thr Leu Ser Asp Ala Asp Arg Lys Gly Gly Asn Ser Ser Val  
 180 185 190

Lys Leu Glu Ala Gly Leu Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys  
 30 195 200 205

Gly Gln Asp Asp Arg Gly Pro Val Val Arg Lys Gln Arg Tyr Val Phe  
 35 210 215 220

Asp Ile Ser Ala Leu Glu Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg  
 225 230 235 240

Ile Leu Arg Lys Lys Pro Ser Asp Thr Ala Lys Pro Ala Ala Pro Gly  
 40 245 250 255

Gly Gly Arg Ala Ala Gln Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg  
 260 265 270

45

- 31 -

Gln Pro Ala Ser Leu Leu Asp Val Arg Ser Val Pro Gly Leu Asp Gly  
 275 280 285

Ser Gly Trp Glu Val Phe Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys  
 5 290 295 300

Asn Ser Ala Gln Leu Cys Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg  
 305 310 315 320

10 Ala Val Asp Leu Arg Gly Leu Gly Phe Asp Arg Ala Ala Arg Gln Val  
 325 330 335

His Glu Lys Ala Leu Phe Leu Val Phe Gly Arg Thr Lys Lys Arg Asp  
 340 345 350

15 Leu Phe Phe Asn Glu Ile Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr  
 355 360 365

Val Tyr Glu Tyr Leu Phe Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu  
 20 370 375 380

Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys  
 385 390 395 400

25 Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp Asp Asp  
 405 410 415

---

Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu  
 420 425 430

30 Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val  
 435 440 445

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr  
 35 450 455 460

Xaa Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp  
 465 470 475 480

40 Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu  
 485 490 495

Ser Cys Gly Cys Arg  
 500

45



<210> 3  
 <211> 2272  
 <212> DNA  
 5 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (128)..(1183)

10  
 <400> 3  
 caaggagcca tgccagctgg acacacactt cttccagggc ctctggcagc caggacagag 60  
 ttgagaccac agctgttgag accctgagcc ctgagtctgt attgctcaag aagggccttc 120

15  
 cccagca atg acc tcc tca ttg ctt ctg gcc ttt ctc ctc ctg gct cca 169  
 Met Thr Ser Ser Leu Leu Leu Ala Phe Leu Leu Leu Ala Pro  
 1 5 10

20 acc aca gtg gcc act ccc aga gct ggc ggt cag tgt cca gca tgt ggg 217  
 Thr Thr Val Ala Thr Pro Arg Ala Gly Gly Gln Cys Pro Ala Cys Gly  
 15 20 25 30

25 ggg ccc acc ttg gaa ctg gag agc cag cgg gag ctg ctt ctt gat ctg 265  
 Gly Pro Thr Leu Glu Leu Glu Ser Gln Arg Glu Leu Leu Leu Asp Leu  
 35 40 45

30 gcc aag aga agc atc ttg gac aag ctg cac ctc acc cag cgc cca aca 313  
 Ala Lys Arg Ser Ile Leu Asp Lys Leu His Leu Thr Gln Arg Pro Thr  
 50 55 60

35 ctg aac cgc cct gtg tcc aga gct gct ttg agg act gca ctg cag cac 361  
 Leu Asn Arg Pro Val Ser Arg Ala Ala Leu Arg Thr Ala Leu Gln His  
 65 70 75

40 ctc cac ggg gtc cca cag ggg gca ctt cta gag gac aac agg gaa cag 409  
 Leu His Gly Val Pro Gln Gly Ala Leu Leu Glu Asp Asn Arg Glu Gln  
 80 85 90

40 gaa tgt gaa atc atc agc ttt gct gag aca ggc ctc tcc acc atc aac 457  
 Glu Cys Glu Ile Ile Ser Phe Ala Glu Thr Gly Leu Ser Thr Ile Asn  
 95 100 105 110

- 33 -

cag act cgt ctt gat ttt cac ttc tcc tct gat aga act gct ggt gac 505  
 Gln Thr Arg Leu Asp Phe His Phe Ser Ser Asp Arg Thr Ala Gly Asp  
 115 120 125

5 agg gag gtc cag cag gcc agt ctc atg ttc ttt gtg cag ctc cct tcc 553  
 Arg Glu Val Gln Gln Ala Ser Leu Met Phe Phe Val Gln Leu Pro Ser  
 130 135 140

10 aat acc act tgg acc ttg aaa gtg aga gtc ctt gtg ctg ggt cca cat 601  
 Asn Thr Thr Trp Thr Leu Lys Val Arg Val Leu Val Leu Gly Pro His  
 145 150 155

aat acc aac ctc acc ttg gct act cag tac ctg ctg gag gtg gat gcc 649  
 Asn Thr Asn Leu Thr Leu Ala Thr Gln Tyr Leu Leu Glu Val Asp Ala  
 15 160 165 170

agt ggc tgg cat caa ctc ccc cta ggg cct gaa gct caa gct gcc tgc 697  
 Ser Gly Trp His Gln Leu Pro Leu Gly Pro Glu Ala Gln Ala Ala Cys  
 175 180 185 190

20 agc cag ggg cac ctg acc ctg gag ctg gta ctt gaa ggc cag gta gcc 745  
 Ser Gln Gly His Leu Thr Leu Glu Leu Val Leu Glu Gly Gln Val Ala  
 195 200 205

---

25 cag agc tca gtc atc ctg ggt gga gct gcc cat agg cct ttt gtg gca 793  
 Gln Ser Ser Val Ile Leu Gly Gly Ala Ala His Arg Pro Phe Val Ala  
 210 215 220

gcc cgg gtg aga gtt ggg ggc aaa cac cag att cac cga cga ggc atc 841  
 30 Ala Arg Val Arg Val Gly Gly Lys His Gln Ile His Arg Arg Gly Ile  
 225 230 235

gac tgc caa gga ggg tcc agg atg tgc tgt cga caa gag ttt ttt gtg 889  
 Asp Cys Gln Gly Gly Ser Arg Met Cys Cys Arg Gln Glu Phe Phe Val  
 35 240 245 250

gac ttc cgt gag att ggc tgg cac gac tgg atc atc cag cct gag ggc 937  
 Asp Phe Arg Glu Ile Gly Trp His Asp Trp Ile Ile Gln Pro Glu Gly  
 255 260 265 270

40 tac gcc atg aac ttc tgc ata ggg cag tgc cca cta cac ata gca ggc 985  
 Tyr Ala Met Asn Phe Cys Ile Gly Gln Cys Pro Leu His Ile Ala Gly  
 275 280 285

- 34 -

atg cct ggt att gct gcc tcc ttt cac act gca gtg ctc aat ctt ctc 1033  
 Met Pro Gly Ile Ala Ala Ser Ph His Thr Ala Val Leu Asn Leu Leu  
 290 295 300

5 aag gcc aac aca gct gca ggc acc act gga ggg ggc tca nnn tgt gta 1081  
 Lys Ala Asn Thr Ala Ala Gly Thr Thr Gly Gly Gly Ser Xaa Cys Val  
 305 310 315

10 ccc acg gcc cgg cgc ccc ctg tct ctg ctc tat tat gac agg gac agc 1129  
 Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu Tyr Tyr Asp Arg Asp Ser  
 320 325 330

15 aac att gtc aag act gac ata cct gac atg gta gta gag gcc tgt ggg 1177  
 Asn Ile Val Lys Thr Asp Ile Pro Asp Met Val Val Glu Ala Cys Gly  
 335 340 345 350

tgc agt tagtctatgt gtggtatggg cagcccaagg ttgcatggga aaacacgccc 1233  
 Cys Ser

20 ctacagaagt gcacttcctt gagaggaggg aatgacctca ttctctgtcc agaattgtgga 1293  
 ctccctcttc ctgagcatct tatggaaatt accccacctt tgacttgaag aaaccttcat 1353  
 ctaaagcaag tcactgtgcc atcttctga ccaactacct ctttcttagg gcatagtcca 1413  
 25 tcccgttagt ccaccccgct agccccaact cagggactca gacccatctc caaccatgag 1473  
 caatgccatc tgggtccag gcaaagacac ccttagctca cctttaatag accccataac 1533

30 ccactatgcc ttctgtcct ttctactcaa tgggtcccaac tccaagatga gttgacacaa 1593  
 ccccttcccc caatttttgt ggatctccag agaggccctt ctttggattc accaaagttt 1653  
 agatcactgc tgcccaaaat agaggcttac ctacccccct ctttgttgtg agccctgtc 1713  
 35 cttcttagtt gtccagggtga actactaaag ctctctttgc ataccttcat ccattttttg 1773  
 tcttctctg cctttctcta tgcccttaag ggggtgactg cctgagctct atcacctgag 1833

40 ctccctgcc ctctggcttc ctgctgaggt cagggcattt cttatccctg ttcctctct 1893  
 gtctaggtgt catggttctg tgtaactgtg gctattctgt gtcctacac tacctggcta 1953  
 ccccttcca tggccccagc tctgctaca ttctgatttt tttttttttt ttttttttga 2013  
 45

aaagttaaaa attccttaatt tttttattcc tgggtaccact accacaattt acagggcaat 2073  
 atacctgatg taatgaaaag aaaaagaaaa agacaaagct acaacagata aaagacctca 2133  
 5 ggaatgtaca tctaattgac actacattgc attaataat agctgcactt tttgcaaact 2193  
 gtggctatga cagtccctgaa caagaagggt ttcctgttta agctgcagta actttttctga 2253  
 ctatggatca tcgttcctt 2272  
 10  
 <210> 4  
 <211> 352  
 <212> PRT  
 15 <213> Homo sapiens  
 <400> 4  
 Met Thr Ser Ser Leu Leu Leu Ala Phe Leu Leu Leu Ala Pro Thr Thr  
 1 5 10 15  
 20 Val Ala Thr Pro Arg Ala Gly Gly Gln Cys Pro Ala Cys Gly Gly Pro  
 20 25 30  
 Thr Leu Glu Leu Glu Ser Gln Arg Glu Leu Leu Leu Asp Leu Ala Lys  
 25 35 40 45  
~~Arg Ser Ile Leu Asp Lys Leu His Leu Thr Gln Arg Pro Thr Leu Asn~~  
 50 55 60  
 30 Arg Pro Val Ser Arg Ala Ala Leu Arg Thr Ala Leu Gln His Leu His  
 65 70 75 80  
 Gly Val Pro Gln Gly Ala Leu Leu Glu Asp Asn Arg Glu Gln Glu Cys  
 85 90 95  
 35 Glu Ile Ile Ser Phe Ala Glu Thr Gly Leu Ser Thr Ile Asn Gln Thr  
 100 105 110  
 Arg Leu Asp Phe His Phe Ser Ser Asp Arg Thr Ala Gly Asp Arg Glu  
 40 115 120 125  
 Val Gln Gln Ala Ser Leu Met Phe Phe Val Gln Leu Pro Ser Asn Thr  
 130 135 140

- 36 -

	Thr Trp Thr Leu Lys Val Arg Val Leu Val Leu Gly Pro His Asn Thr	
	145	150 155 160
5	Asn Leu Thr Leu Ala Thr Gln Tyr Leu Leu Glu Val Asp Ala Ser Gly	
	165	170 175
	Trp His Gln Leu Pro Leu Gly Pro Glu Ala Gln Ala Ala Cys Ser Gln	
	180	185 190
10	Gly His Leu Thr Leu Glu Leu Val Leu Glu Gly Gln Val Ala Gln Ser	
	195	200 205
	Ser Val Ile Leu Gly Gly Ala Ala His Arg Pro Phe Val Ala Ala Arg	
	210	215 220
15	Val Arg Val Gly Gly Lys His Gln Ile His Arg Arg Gly Ile Asp Cys	
	225	230 235 240
	Gln Gly Gly Ser Arg Met Cys Cys Arg Gln Glu Phe Phe Val Asp Phe	
20	245	250 255
	Arg Glu Ile Gly Trp His Asp Trp Ile Ile Gln Pro Glu Gly Tyr Ala	
	260	265 270
25	Met Asn Phe Cys Ile Gly Gln Cys Pro Leu His Ile Ala Gly Met Pro	
	275	280 285
	Gly Ile Ala Ala Ser Phe His Thr Ala Val Leu Asn Leu Leu Lys Ala	
	290	295 300
30	Asn Thr Ala Ala Gly Thr Thr Gly Gly Gly Ser Xaa Cys Val Pro Thr	
	305	310 315 320
	Ala Arg Arg Pro Leu Ser Leu Leu Tyr Tyr Asp Arg Asp Ser Asn Ile	
35	325	330 335
	Val Lys Thr Asp Ile Pro Asp Met Val Val Glu Ala Cys Gly Cys Ser	
	340	345 350
40		
	<210> 5	
	<211> 30	
	<212> DNA	
45	<213> Artificial Sequence	

- 37 -

&lt;220&gt;

<223> Description of Artificial Sequence: mutagenic  
primer

5 &lt;400&gt; 5

ccacaccacc caccgcctgt gtgccacgc

30

10 &lt;210&gt; 6

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

15 &lt;220&gt;

<223> Description of Artificial Sequence: mutagenic  
primer

&lt;400&gt; 6

20 gcgtgggcac acaggcgtg ggtggtgtgg

30

&lt;210&gt; 7

25 &lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

30 <223> Description of Artificial Sequence: mutagenic  
primer

&lt;400&gt; 7

35 ctggaggggg ctcagcctgt gtaccacgg

30

&lt;210&gt; 8

&lt;211&gt; 30

40 &lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

45 <223> Description of Artificial Sequence: mutagenic  
primer

<400> 8

ccgtgggtac acaggctgag cccctccag

30

## Claims

1. Protein selected from the members of the TGF- $\beta$  superfamily,  
characterized in that the protein is necessarily monomeric due to  
substitution or deletion of a cysteine which is responsible for dimer  
formation.
2. Protein according to claim 1,  
characterized in that the protein is a mature protein or a biologically  
active part or variant thereof.
3. Protein according to any one of the preceeding claims,  
characterized in that the protein contains at least the 7 cysteine  
region characteristic for the TGF- $\beta$  protein superfamily.
4. Protein according to claim 3,  
characterized in that it contains a consensus sequence according to  
Formula I: C(Y)<sub>25-29</sub>CYYC(Y)<sub>25-35</sub>XC(Y)<sub>27-34</sub>CYC or  
Formula II: C(Y)<sub>28</sub>CYYC(Y)<sub>30-32</sub>XC(Y)<sub>31</sub>CYC, wherein C denotes  
cysteine, Y denotes any amino acid and X denotes any amino acid  
except cysteine.
5. Protein according to any one of claims 1 to 4,  
characterized in that the protein is a morphogenetic protein.
6. Protein according to any one of the preceeding claims,  
characterized in that the proteins belongs to the TGF- $\beta$ , BMP, GDF,  
activin or GDNF family.
7. Protein according to claim 6,  
characterized in that the protein is MP52 (GDF5) or a biologically  
active part or variant thereof.



8. Protein according to any one of the preceeding claims,  
characterized in that it comprises the amino acid sequence according  
to SEQ.ID.NO.2 or a part thereof.
- 5 9. Protein according to claim 6,  
characterized in that the protein is MP121 or a biologically active part  
or variant thereof.
- 10 10. Protein according to claim 9,  
characterized in that it comprises the amino acid sequence according  
to SEQ.ID.NO.4 or a part thereof.
- 15 11. Protein according to any one of claims 1 to 10,  
characterized in that the cysteine residue is substituted by an amino  
acid selected from the group of alanine, serine, threonine, leucine,  
isoleucine, glycine and valine.
- 20 12. Protein according to any one of claims 1 to 11,  
characterized in that it contains additional amino acids that facilitate  
or mediate the transfer and localization of the protein in a certain  
tissue.
- 25 13. Nucleic acid,  
characterized in that it encodes a protein according to any one of  
claims 1 to 12.
14. Nucleic acid according to claim 13,  
characterized in that it is a DNA.
- 30 15. Nucleic acid according to claim 13 or 14,  
characterized in that it contains a sequence as shown in  
SEQ.ID.NO.1 or a fragment thereof.

16. Nucleic acid according to claim 13 or 14,  
characterized in that it contains a sequence as shown in  
SEQ.ID.NO.3 or a fragment thereof.
- 5 17. Nucleic acid according to any one of claims 13 to 16,  
characterized in that it further contains suitable expression control  
sequences facilitating and/or driving expression of the encoded  
protein.
- 10 18. Expression vector,  
characterized in that it contains a nucleic acid according to any one  
of claims 13 to 17 in a suitable vector system.
- 15 19. Expression vector according to claim 18,  
characterized in that the vector system is suitable for prokaryotic  
expression.
20. Host cell,  
characterized in that it contains a nucleic acid according to any one  
20 of claims 13 to 17 or an expression vector according to claims 18 or  
19 and upon expression of said nucleic acid or vector is able to  
produce a monomeric protein according to any one of claims 1 to 12.
21. Host cell according to claim 20,  
25 characterized in that it is a prokaryotic host cell.
22. Host cell according to claim 20,  
characterized in that it is an embryonal cell.
- 30 23. Pharmaceutical composition,  
characterized in that it contains at least one protein according to any  
one of claims 1 to 12 or at least one nucleic acid according to any

one of claims 13 to 17, at least one expression vector according to any one of claims 18 or 19 or at least one host cell according to claim 20 or 22.

- 5      24.    Pharmaceutical composition according to claim 23,  
characterized in that the protein and/or nucleic acid are contained in  
and/or on a biocompatible matrix material.
- 10      25.    Pharmaceutical composition according to claim 24,  
characterized in that the matrix material is biodegradable.
26.    Pharmaceutical composition according to claims 24 or 25,  
characterized in that the matrix material is itself osteogenically active.
- 15      27.    Pharmaceutical composition according to any one of claims 23 to 26,  
for the prevention or therapy of diseases for which also the dimeric  
form of the protein would be indicated.
- 20      28.    Pharmaceutical composition according to claim 27,  
for prevention or therapy of diseases associated with bone and/or  
cartilage damage or affecting bone and/or cartilage disease or  
situations in which cartilage and/or bone growth is desirable or for  
spinal fusion.
- 25      29.    Pharmaceutical composition according to claim 27,  
for prevention or therapy of damaged or diseased tissue associated  
with connective tissue including tendon and/or ligament, periodontal  
or dental tissue including dental implants, neural tissue including CNS  
tissue and neuropathological situations, tissue of the sensory system,  
30      liver, pancreas, cardiac, blood vessel, renal, uterine and thyroid  
tissue, skin, mucous membranes, endothelium, epithelium, for  
promotion or induction of nerve growth, tissue regeneration,

angiogenesis, wound healing including ulcers, burns, injuries or skin grafts, induction of proliferation of progenitor cells or bone marrow cells, for maintenance of a state of proliferation or differentiation, for treatment or preservation of tissue or cells for organ or tissue transplantation, for integrity of gastrointestinal lining, for treatment of disturbances in fertility, contraception or pregnancy.

30. Pharmaceutical composition according to any one of claims 23 to 29 for surgical local application, topical or systemic application.

31. Pharmaceutical composition according to any one of claims 23 to 30 characterized in that it further contains pharmacologically acceptable auxiliary substances.

32. Pharmaceutical composition according to any one of claims 30 or 31, characterized in that the composition is injectable.

33. Pharmaceutical composition according to anyone of claims 30 to 32, ~~characterized in that it is contained in a vehicle that allows to direct and release the composition to a determined site of action.~~

34. Pharmaceutical composition according to claim 33, characterized in that the vehicle is selected from liposomes, nanospheres, larger implantable containers and micropumps.

35. Use of a pharmaceutical composition according to any one of claims 23 to 34 for the prevention or treatment of any indications of the dimeric form of the protein.

Fig. 1 A

Name: MP52, dimeric form

Formula

 $C_{1184}H_{1844}N_{330}O_{350}S_{22}$ 

Molecular weight

26994 Dalton

Amino acid composition

238 amino acids

Disulfide bond

7 bonds

## Primary structure

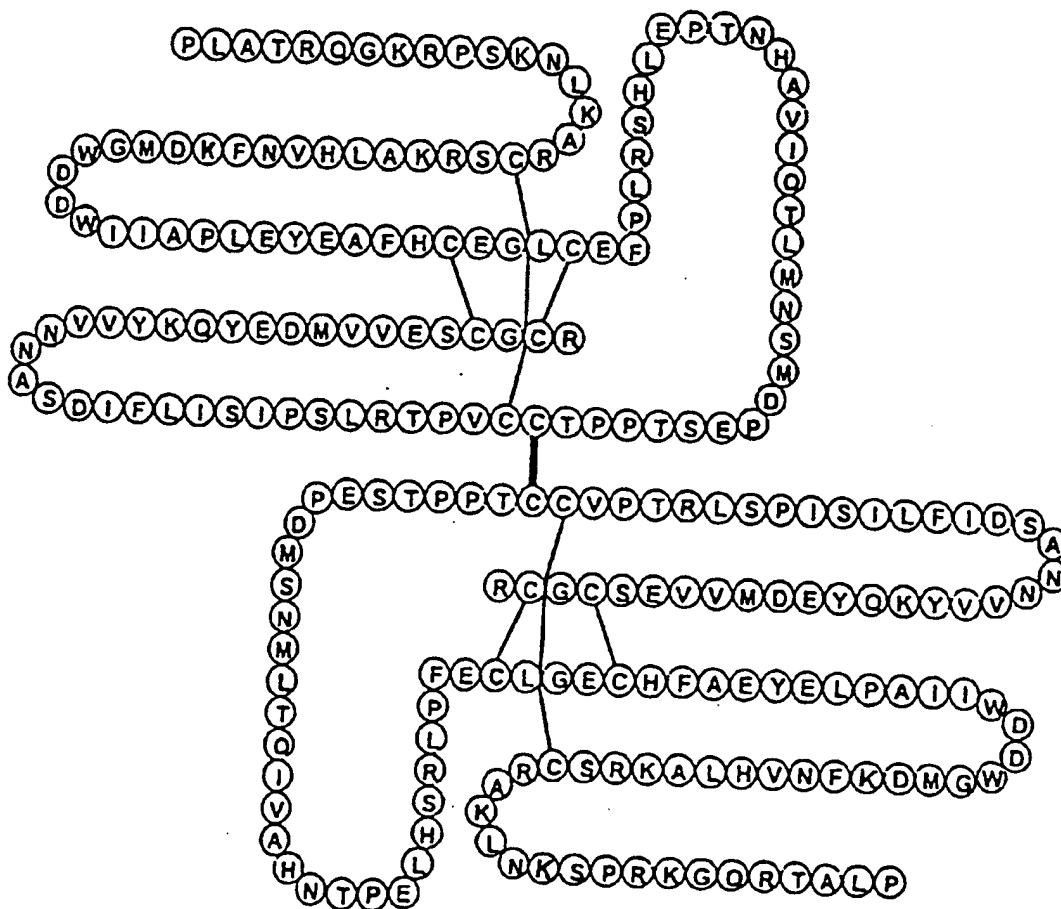
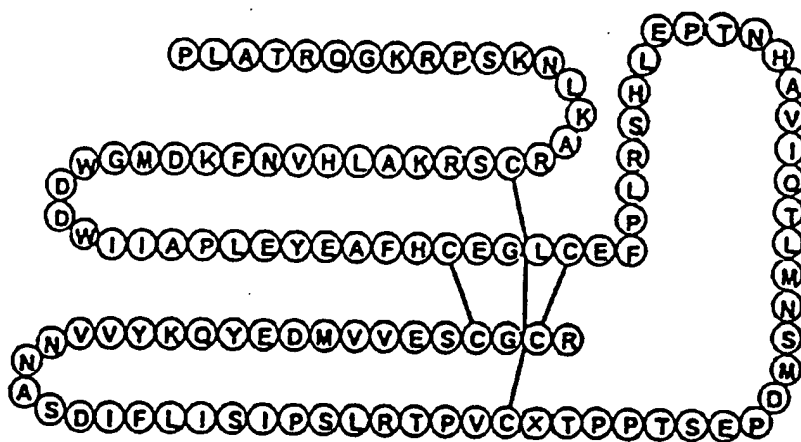


Fig. 1 B



## SEQUENCE LISTING

<110> HyGene AG

<120> Monomeric Protein of the TGF-beta Family

<130> 20780pwomd

<140>

<141>

<150> EP99115613

<151> 1999-08-06

<160> 8

<170> PatentIn Ver. 2.1

<210> 1

<211> 2703

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (640)..(2142)

<400> 1

ccatggcctc gaaagggcag cggtgatttt tttcacataa atatatcgca cttaaagtga 60  
tttagacagc atgacatcag agagtaatta aattggtttg ggttggaatt ccgtttccaa 120  
ttcctgagtt caggtttgta aaagattttt ctgagcacct gcaggcctgt gagtgtgtgt 180  
gtgtgtgtgt gtgtgtgtgt gtgtgtgtga agtattttca ctggaaagga ttcaaaacta 240  
gggggaaaaa aaaactggag cacacaggca gcattacgcc attcttcctt ctggaaaaa 300  
tccttcagcc ttatacaagc ctcttcaag ccctcagtca gttgtgcagg agaaaggggg 360  
cggttggett tctcctttca agaacgagtt attttcagct gctgactgga gacgggtgcac 420  
gtctggatac gagagcattt ccactatggg actggataca aacacacacc cggcagactt 480  
caagagtctc agactgagga gaaagccttt ccttctgctg ctactgctgc tgccgctgct 540  
tttgaaagtc cactcctttc atggtttttc ctgccaaacc agaggcacct ttgctgctgc 600

cgctgttctc tttggtgtca ttcagcggct ggccagagg atg aga ctc ccc aaa	654
Met Arg Leu Pro Lys	
1 5	
ctc ctc act ttc ttg ctt tgg tac ctg gct tgg ctg gac ctg gaa ttc	702
Leu Leu Thr Phe Leu Leu Trp Tyr Leu Ala Trp Leu Asp Leu Glu Phe	
10 15 20	
atc tgc act gtg ttg ggt gcc cct gac ttg ggc cag aga ccc cag ggg	750
Ile Cys Thr Val Leu Gly Ala Pro Asp Leu Gly Gln Arg Pro Gln Gly	
25 30 35	
acc agg cca gga ttg gcc aaa gca gag gcc aag gag agg ccc ccc ctg	798
Thr Arg Pro Gly Leu Ala Lys Ala Glu Ala Lys Glu Arg Pro Pro Leu	
40 45 50	
gcc cgg aac gtc ttc agg cca ggg ggt cac agc tat ggt ggg ggg gcc	846
Ala Arg Asn Val Phe Arg Pro Gly Gly His Ser Tyr Gly Gly Gly Ala	
55 60 65	
acc aat gcc aat gcc agg gca aag gga ggc acc ggg cag aca gga ggc	894
Thr Asn Ala Asn Ala Arg Ala Lys Gly Gly Thr Gly Gln Thr Gly Gly	
70 75 80 85	
ctg aca cag ccc aag aag gat gaa ccc aaa aag ctg ccc ccc aga ccg	942
Leu Thr Gln Pro Lys Lys Asp Glu Pro Lys Lys Leu Pro Pro Arg Pro	
90 95 100	
ggc gcc cct gaa ccc aag cca gga cac cct ccc caa aca agg cag gct	990
Gly Gly Pro Glu Pro Lys Pro Gly His Pro Pro Gln Thr Arg Gln Ala	
105 110 115	
aca gcc cgg act gtg acc cca aaa gga cag ctt ccc gga ggc aag gca	1038
Thr Ala Arg Thr Val Thr Pro Lys Gly Gln Leu Pro Gly Gly Lys Ala	
120 125 130	
ccc cca aaa gca gga tct gtc ccc agc tcc ttc ctg ctg aag aag gcc	1086
Pro Pro Lys Ala Gly Ser Val Pro Ser Ser Phe Leu Leu Lys Lys Ala	
135 140 145	
agg gag ccc ggg ccc cca cga gag ccc aag gag ccg ttt cgc cca ccc	1134
Arg Glu Pro Gly Pro Pro Arg Glu Pro Lys Glu Pro Phe Arg Pro Pro	
150 155 160 165	
ccc atc aca ccc cac gag tac atg ctc tcg ctg tac agg acg ctg tcc	1182
Pro Ile Thr Pro His Glu Tyr Met Leu Ser Leu Tyr Arg Thr Leu Ser	
170 175 180	



gat gct gac aga aag gga ggc aac agc agc gtg aag ttg gag gct ggc	1230
Asp Ala Asp Arg Lys Gly Gly Asn Ser Ser Val Lys Leu Glu Ala Gly	
185 190 195	
ctg gcc aac acc atc acc agc ttt att gac aaa ggg caa gat gac cga	1278
Leu Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys Gly Gln Asp Asp Arg	
200 205 210	
ggc ccc gtg gtc agg aag cag agg tac gtg ttt gac att agt gcc ctg	1326
Gly Pro Val Val Arg Lys Gln Arg Tyr Val Phe Asp Ile Ser Ala Leu	
215 220 225	
gag aag gat ggg ctg ctg ggg gcc gag ctg cgg atc ttg cgg aag aag	1374
Glu Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg Ile Leu Arg Lys Lys	
230 235 240 245	
ccc tcg gac acg gcc aag cca gcg gcc ccc gga ggc ggg cgg gct gcc	1422
Pro Ser Asp Thr Ala Lys Pro Ala Ala Pro Gly Gly Gly Arg Ala Ala	
250 255 260	
cag ctg aag ctg tcc agc tgc ccc agc ggc cgg cag ccg gcc tcc ttg	1470
Gln Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg Gln Pro Ala Ser Leu	
265 270 275	
ctg gat gtg cgc tcc gtg cca ggc ctg gac gga tct ggc tgg gag gtg	1518
Leu Asp Val Arg Ser Val Pro Gly Leu Asp Gly Ser Gly Trp Glu Val	
280 285 290	
ttc gac atc tgg aag ctc ttc cga aac ttt aag aac tcg gcc cag ctg	1566
Phe Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys Asn Ser Ala Gln Leu	
295 300 305	
tgc ctg gag ctg gag gcc tgg gaa cgg ggc agg gcc gtg gac ctc cgt	1614
Cys Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg Ala Val Asp Leu Arg	
310 315 320 325	
ggc ctg ggc ttc gac cgc gcc gcc cgg cag gtc cac gag aag gcc ctg	1662
Gly Leu Gly Phe Asp Arg Ala Ala Arg Gln Val His Glu Lys Ala Leu	
330 335 340	
ttc ctg gtg ttt ggc cgc acc aag aaa cgg gac ctg ttc ttt aat gag	1710
Phe Leu Val Phe Gly Arg Thr Lys Lys Arg Asp Leu Phe Phe Asn Glu	
345 350 355	
att aag gcc cgc tct ggc cag gac gat aag acc gtg tat gag tac ctg	1758
Ile Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr Val Tyr Glu Tyr Leu	
360 365 370	

ttc agc cag cgg cga aaa cgg cgg gcc cca ctg gcc act cgc cag ggc 1806  
 Phe Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu Ala Thr Arg Gln Gly  
 375 380 385

aag cga ccc agc aag aac ctt aag gct cgc tgc agt cgg aag gca ctg 1854  
 Lys Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys Ser Arg Lys Ala Leu  
 390 395 400 405

cat gtc aac ttc aag gac atg ggc tgg gac gac tgg atc atc gca ccc 1902  
 His Val Asn Phe Lys Asp Met Gly Trp Asp Asp Trp Ile Ile Ala Pro  
 410 415 420

ctt gag tac gag gct ttc cac tgc gag ggg ctg tgc gag ttc cca ttg 1950  
 Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu  
 425 430 435

cgc tcc cac ctg gag ccc acg aat cat gca gtc atc cag acc ctg atg 1998  
 Arg Ser His Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met  
 440 445 450

aac tcc atg gac ccc gag tcc aca cca ccc acc nnn tgt gtg ccc acg 2046  
 Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr Xaa Cys Val Pro Thr  
 455 460 465

cgg ctg agt ccc atc agc atc ctc ttc att gac tct gcc aac aac gtg 2094  
 Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val  
 470 475 480 485

---

gtg tat aag cag tat gag gac atg gtc gtg gag tcg tgt ggc tgc agg 2142  
 Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys Arg  
 490 495 500

tagcagcact ggccctctgt cttctctgggt ggcacatccc aagagccctt tcctgcactc 2202

ctggaatcac agaggggtca ggaagctgtg gcaggagcat ctacacagct tgggtgaaag 2262

gggattccaa taagcttgct cgctctctga gtgtgacttg ggctaaaggc ccctttttat 2322

ccacaagtgc ccctggctga ggattgctgc ccgtctgctg atgtgaccag tggcaggcac 2382

aggtccaggg agacagactc tgaatgggac tgagtcccag gaaacagtgc tttccgatga 2442

gactcagccc accatttctc ctcacctggg ccttctcagc ctctggactc tcctaagcac 2502

ctctcaggag agccacaggt gccactgcct cctcaaatca catttggtgcc tgggtgacttc 2562

ctgtccctgg gacagttgag aagctgactg ggcaagagtg ggagagaaga ggagagggct 2622

tggatagagt tgaggagtgt gaggtgttta gactgttaga tttaaatgta tattgatgag 2682

ataaaaagca aaactgtgcc t

2703

<210> 2

<211> 501

<212> PRT

<213> Homo sapiens

<400> 2

Met Arg Leu Pro Lys Leu Leu Thr Phe Leu Leu Trp Tyr Leu Ala Trp  
1 5 10 15

Leu Asp Leu Glu Phe Ile Cys Thr Val Leu Gly Ala Pro Asp Leu Gly  
20 25 30

Gln Arg Pro Gln Gly Thr Arg Pro Gly Leu Ala Lys Ala Glu Ala Lys  
35 40 45

Glu Arg Pro Pro Leu Ala Arg Asn Val Phe Arg Pro Gly Gly His Ser  
50 55 60

Tyr Gly Gly Gly Ala Thr Asn Ala Asn Ala Arg Ala Lys Gly Gly Thr  
65 70 75 80

Gly Gln Thr Gly Gly Leu Thr Gln Pro Lys Lys Asp Glu Pro Lys Lys  
85 90 95

Leu Pro Pro Arg Pro Gly Gly Pro Glu Pro Lys Pro Gly His Pro Pro  
100 105 110

Gln Thr Arg Gln Ala Thr Ala Arg Thr Val Thr Pro Lys Gly Gln Leu  
115 120 125

Pro Gly Gly Lys Ala Pro Pro Lys Ala Gly Ser Val Pro Ser Ser Phe  
130 135 140

Leu Leu Lys Lys Ala Arg Glu Pro Gly Pro Pro Arg Glu Pro Lys Glu  
145 150 155 160

Pro Phe Arg Pro Pro Pro Ile Thr Pro His Glu Tyr Met Leu Ser Leu  
165 170 175

Tyr Arg Thr Leu Ser Asp Ala Asp Arg Lys Gly Gly Asn Ser Ser Val  
180 185 190

Lys Leu Glu Ala Gly Leu Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys

195	200	205
Gly Gln Asp Asp Arg Gly Pro Val Val Arg Lys Gln Arg Tyr Val Phe		
210	215	220
Asp Ile Ser Ala Leu Glu Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg		
225	230	235 240
Ile Leu Arg Lys Lys Pro Ser Asp Thr Ala Lys Pro Ala Ala Pro Gly		
	245	250 255
Gly Gly Arg Ala Ala Gln Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg		
	260	265 270
Gln Pro Ala Ser Leu Leu Asp Val Arg Ser Val Pro Gly Leu Asp Gly		
	275	280 285
Ser Gly Trp Glu Val Phe Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys		
	290	295 300
Asn Ser Ala Gln Leu Cys Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg		
305	310	315 320
Ala Val Asp Leu Arg Gly Leu Gly Phe Asp Arg Ala Ala Arg Gln Val		
	325	330 335
His Glu Lys Ala Leu Phe Leu Val Phe Gly Arg Thr Lys Lys Arg Asp		
	340	345 350
Leu Phe Phe Asn Glu Ile Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr		
	355	360 365
Val Tyr Glu Tyr Leu Phe Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu		
	370	375 380
Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys		
385	390	395 400
Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp Asp Asp		
	405	410 415
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu		
	420	425 430
Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val		
	435	440 445
Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr		

450                      455                      460  
 Xaa Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp  
 465                      470                      475                      480  
 Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu  
                          485                      490                      495  
 Ser Cys Gly Cys Arg  
                          500

<210> 3  
 <211> 2272  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (128)..(1183)

<400> 3  
 caaggagcca tgccagctgg acacacactt cttccagggc ctctggcagc caggacagag 60  
 ttgagaccac agctgttgag accctgagcc ctgagctctgt attgctcaag aagggccttc 120  
 cccagca atg acc tcc tca ttg ctt ctg gcc ttt ctc ctc ctg gct cca 169  
          Met Thr Ser Ser Leu Leu Leu Ala Phe Leu Leu Leu Ala Pro  
               1                      5                      10  
 acc aca gtg gcc act ccc aga gct ggc ggt cag tgt cca gca tgt ggg 217  
 Thr Thr Val Ala Thr Pro Arg Ala Gly Gly Gln Cys Pro Ala Cys Gly  
   15                      20                      25                      30  
 ggg ccc acc ttg gaa ctg gag agc cag cgg gag ctg ctt ctt gat ctg 265  
 Gly Pro Thr Leu Glu Leu Glu Ser Gln Arg Glu Leu Leu Leu Asp Leu  
                          35                      40                      45  
 gcc aag aga agc atc ttg gac aag ctg cac ctc acc cag cgc cca aca 313  
 Ala Lys Arg Ser Ile Leu Asp Lys Leu His Leu Thr Gln Arg Pro Thr  
                          50                      55                      60  
 ctg aac cgc cct gtg tcc aga gct gct ttg agg act gca ctg cag cac 361  
 Leu Asn Arg Pro Val Ser Arg Ala Ala Leu Arg Thr Ala Leu Gln His  
                          65                      70                      75  
 ctc cac ggg gtc cca cag ggg gca ctt cta gag gac aac agg gaa cag 409

Leu His Gly Val Pro Gln Gly Ala Leu Leu Glu Asp Asn Arg Glu Gln	
80 85 90	
gaa tgt gaa atc atc agc ttt gct gag aca ggc ctc tcc acc atc aac	457
Glu Cys Glu Ile Ile Ser Phe Ala Glu Thr Gly Leu Ser Thr Ile Asn	
95 100 105 110	
cag act cgt ctt gat ttt cac ttc tcc tct gat aga act gct ggt gac	505
Gln Thr Arg Leu Asp Phe His Phe Ser Ser Asp Arg Thr Ala Gly Asp	
115 120 125	
agg gag gtc cag cag gcc agt ctc atg ttc ttt gtg cag ctc cct tcc	553
Arg Glu Val Gln Gln Ala Ser Leu Met Phe Phe Val Gln Leu Pro Ser	
130 135 140	
aat acc act tgg acc ttg aaa gtg aga gtc ctt gtg ctg ggt cca cat	601
Asn Thr Thr Trp Thr Leu Lys Val Arg Val Leu Val Leu Gly Pro His	
145 150 155	
aat acc aac ctc acc ttg gct act cag tac ctg ctg gag gtg gat gcc	649
Asn Thr Asn Leu Thr Leu Ala Thr Gln Tyr Leu Leu Glu Val Asp Ala	
160 165 170	
agt ggc tgg cat caa ctc ccc cta ggg cct gaa gct caa gct gcc tgc	697
Ser Gly Trp His Gln Leu Pro Leu Gly Pro Glu Ala Gln Ala Ala Cys	
175 180 185 190	
agc cag ggg cac ctg acc ctg gag ctg gta ctt gaa ggc cag gta gcc	745
Ser Gln Gly His Leu Thr Leu Glu Leu Val Leu Glu Gly Gln Val Ala	
195 200 205	
cag agc tca gtc atc ctg ggt gga gct gcc cat agg cct ttt gtg gca	793
Gln Ser Ser Val Ile Leu Gly Gly Ala Ala His Arg Pro Phe Val Ala	
210 215 220	
gcc cgg gtg aga gtt ggg ggc aaa cac cag att cac cga cga ggc atc	841
Ala Arg Val Arg Val Gly Gly Lys His Gln Ile His Arg Arg Gly Ile	
225 230 235	
gac tgc caa gga ggg tcc agg atg tgc tgt cga caa gag ttt ttt gtg	889
Asp Cys Gln Gly Gly Ser Arg Met Cys Cys Arg Gln Glu Phe Phe Val	
240 245 250	
gac ttc cgt gag att ggc tgg cac gac tgg atc atc cag cct gag ggc	937
Asp Phe Arg Glu Ile Gly Trp His Asp Trp Ile Ile Gln Pro Glu Gly	
255 260 265 270	
tac gcc atg aac ttc tgc ata ggg cag tgc cca cta cac ata gca ggc	985

Tyr Ala Met Asn Phe Cys Il Gly Gln Cys Pro Leu His Ile Ala Gly  
 275 280 285

atg cct ggt att gct gcc tcc ttt cac act gca gtg ctc aat ctt ctc 1033  
 Met Pro Gly Ile Ala Ala Ser Phe His Thr Ala Val Leu Asn Leu Leu  
 290 295 300

aag gcc aac aca gct gca ggc acc act gga ggg ggc tca nnn tgt gta 1081  
 Lys Ala Asn Thr Ala Ala Gly Thr Thr Gly Gly Gly Ser Xaa Cys Val  
 305 310 315

ccc acg gcc cgg cgc ccc ctg tct ctg ctc tat tat gac agg gac agc 1129  
 Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu Tyr Tyr Asp Arg Asp Ser  
 320 325 330

aac att gtc aag act gac ata cct gac atg gta gta gag gcc tgt ggg 1177  
 Asn Ile Val Lys Thr Asp Ile Pro Asp Met Val Val Glu Ala Cys Gly  
 335 340 345 350

tgc agt tagtctatgt gtggtatggg cagcccaagg ttgcatggga aaacacgccc 1233  
 Cys Ser

ctacagaagt gcacttcctt gagaggaggg aatgacctca ttctctgtcc agaattgtga 1293

ctccctcttc ctgagcatct tatggaaatt accccacctt tgacttgaag aaaccttcat 1353

ctaaagcaag tcaactgtgcc atcttcctga ccactaccct ctttcttagg gcatagtcca 1413

tcccgtagt ccattcccgt agccccactc cagggactca gacccatctc caaccatgag 1473

caatgccatc tgggtcccag gcaaagacac ccttagctca cctttaatag accccataac 1533

ccactatgcc ttctgtcct ttctactcaa tgggtcccccac tccaagatga gttgacacaa 1593

ccccttcccc caatttttgt ggatctccag agaggccctt ctttggattc accaaagttt 1653

agatcactgc tgcccaaat agaggcttac ctacccccct ctttgttgtg agcccctgtc 1713

cttcttagtt gtccaggtga actactaaag ctctctttgc ataccttcat ccattttttg 1773

tccttctctg cctttctcta tgcccttaag gggtgacttg cctgagctct atcacctgag 1833

ctcccctgcc ctctggcttc ctgctgaggt cagggcattt cttatccctg ttccctctct 1893

gtctaggtgt catggttctg tgtaactgtg gctattctgt gtccctacac tacctggeta 1953

ccccttcca tggccccagc tctgcctaca ttctgatttt tttttttttt ttttttttga 2013

aaagttaaaa attccttaat tttttattcc tggtagcact accacaattt acagggaat 2073  
 atacctgatg taatgaaaag aaaaagaaaa agacaaagct acaacagata aaagacctca 2133  
 ggaatgtaca tctaattgac actacattgc attaatcaat agctgcactt ttgcaaact 2193  
 gtggctatga cagtcctgaa caagaagggt ttcctgttta agctgcagta acttttctga 2253  
 cratggatca tcgttcctt 2272

&lt;210&gt; 4

&lt;211&gt; 352

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

Met Thr Ser Ser Leu Leu Leu Ala Phe Leu Leu Leu Ala Pro Thr Thr  
 1 5 10 15

Val Ala Thr Pro Arg Ala Gly Gly Gln Cys Pro Ala Cys Gly Gly Pro  
 20 25 30

Thr Leu Glu Leu Glu Ser Gln Arg Glu Leu Leu Leu Asp Leu Ala Lys  
 35 40 45

Arg Ser Ile Leu Asp Lys Leu His Leu Thr Gln Arg Pro Thr Leu Asn  
 50 55 60

Arg Pro Val Ser Arg Ala Ala Leu Arg Thr Ala Leu Gln His Leu His  
 65 70 75 80

Gly Val Pro Gln Gly Ala Leu Leu Glu Asp Asn Arg Glu Gln Glu Cys  
 85 90 95

Glu Ile Ile Ser Phe Ala Glu Thr Gly Leu Ser Thr Ile Asn Gln Thr  
 100 105 110

Arg Leu Asp Phe His Phe Ser Ser Asp Arg Thr Ala Gly Asp Arg Glu  
 115 120 125

Val Gln Gln Ala Ser Leu Met Phe Phe Val Gln Leu Pro Ser Asn Thr  
 130 135 140

Thr Trp Thr Leu Lys Val Arg Val Leu Val Leu Gly Pro His Asn Thr  
 145 150 155 160

Asn Leu Thr Leu Ala Thr Gln Tyr Leu Leu Glu Val Asp Ala Ser Gly



	165		170		175
Trp His Gln Leu Pro Leu Gly Pro Glu Ala Gln Ala Ala Cys Ser Gln					
	180		185		190
Gly His Leu Thr Leu Glu Leu Val Leu Glu Gly Gln Val Ala Gln Ser					
	195		200		205
Ser Val Ile Leu Gly Gly Ala Ala His Arg Pro Phe Val Ala Ala Arg					
	210		215		220
Val Arg Val Gly Gly Lys His Gln Ile His Arg Arg Gly Ile Asp Cys					
	225		230		235
					240
Gln Gly Gly Ser Arg Met Cys Cys Arg Gln Glu Phe Phe Val Asp Phe					
	245		250		255
Arg Glu Ile Gly Trp His Asp Trp Ile Ile Gln Pro Glu Gly Tyr Ala					
	260		265		270
Met Asn Phe Cys Ile Gly Gln Cys Pro Leu His Ile Ala Gly Met Pro					
	275		280		285
Gly Ile Ala Ala Ser Phe His Thr Ala Val Leu Asn Leu Leu Lys Ala					
	290		295		300
Asn Thr Ala Ala Gly Thr Thr Gly Gly Gly Ser Xaa Cys Val Pro Thr					
	305		310		315
					320
Ala Arg Arg Pro Leu Ser Leu Leu Tyr Tyr Asp Arg Asp Ser Asn Ile					
	325		330		335
Val Lys Thr Asp Ile Pro Asp Met Val Val Glu Ala Cys Gly Cys Ser					
	340		345		350

&lt;210&gt; 5

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: mutagenic  
primer

&lt;400&gt; 5

ccacaccacc caccgctgt gtgccacgc

30

<210> 6  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: mutagenic  
primer

<400> 6  
gcgtgggcac acaggcggtg ggtggtgtgg 30

<210> 7  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: mutagenic  
primer

<400> 7  
ctggaggggg ctcagcctgt gtacccacgg 30

---

<210> 8  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: mutagenic  
primer

<400> 8  
ccgtgggtac acaggctgag cccctccag 30

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 00/07600

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/495 A61K38/18

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, WPI Data, EMBASE, MEDLINE, CHEM ABS Data, BIOSIS, STRAND

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HÜSKEN-HINDI, PETRA ET AL: "Monomeric activin A retains high receptor binding affinity but exhibits low biological activity." JOURNAL OF BIOLOGICAL CHEMISTRY, (1994) VOL. 269, NO. 30, PP. 19380-19384, 24 July 1994 (1994-07-24), XP002111996	1-6,11, 13,14, 17,18,20
Y	the whole document	7-10,12, 15,16, 19,21-35
	— -/-	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*A\* document member of the same patent family

Date of the actual completion of the international search

14 December 2000

Date of mailing of the international search report

21/12/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Hix, R

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 00/07600

## C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AMATAYAKUL-CHANTLER, SUPAVADEE (1) ET AL: "(Ser-77)Transforming growth factor-beta-1: Selective biological activity and receptor binding in mink lung epithelial cells." JOURNAL OF BIOLOGICAL CHEMISTRY, (1994) VOL. 269, NO. 44, PP. 27687-27691, 4 November 1994 (1994-11-04), XP002111995	1-6,11, 13,14, 17-22
Y	the whole document	7-10,12, 15,16, 23-35
X	BRUNNER, AMY M. ET AL: "Site-directed mutagenesis of glycosylation sites in the transforming growth factor-beta-1 ( TGF - beta -1) and TGF - beta -2 (414) precursors and of cysteine residues within mature TGF - beta -1: Effects on secretion and bioactivity." MOLECULAR ENDOCRINOLOGY, (1992) VOL. 6, NO. 10, PP. 1691-1700., XP000863262	1-6,11, 13,14, 17,18,20
Y	the whole document	7-10,12, 15,16, 19,21-35
X	BRUNNER A M ET AL: "SITE-DIRECTED MUTAGENESIS OF CYSTEINE RESIDUES IN THE PRO REGION OF THE TRANSFORMING GROWTH FACTOR BETA-1 PRECURSOR EXPRESSION AND CHARACTERIZATION OF MUTANT PROTEINS." J BIOL CHEM, (1989) 264 (23), 13660-13664., XP000857484	1,3-6
Y	the whole document	
Y	WO 94 17099 A (CELTRIX PHARMA) 4 August 1994 (1994-08-04) the whole document	1-35
Y	WO 95 04819 A (BIOPH BIOTECH ENTW PHARM GMBH) 16 February 1995 (1995-02-16) the whole document	1-35
Y	WO 96 01316 A (BIOPH BIOTECH ENTW PHARM GMBH ;HOETTEN GERTRUD (DE); NEIDHARDT HEL) 18 January 1996 (1996-01-18) the whole document	1-35
Y	WO 95 16035 A (GENETICS INST ;HARVARD COLLEGE (US)) 15 June 1995 (1995-06-15) the whole document	1-35
	— -/-	

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 00/07600

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 915 168 A (HOECHST MARION ROUSSEL LTD) 12 May 1999 (1999-05-12) the whole document	1-35
Y	WO 97 03188 A (BIOPH BIOTECH ENTW PHARM GMBH ;HOETTEN GERTRUD (DE); POHL JENS (DE)) 30 January 1997 (1997-01-30) cited in the application the whole document	1-35
A	DAOPIN S ET AL: "CRYSTAL STRUCTURE OF TRANSFORMING GROWTH FACTOR-BETA2: UNUSUAL FOLD FOR THE SUPERFAMILY" SCIENCE,US,AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, vol. 257, page 369-373 XP000857489 ISSN: 0036-8075 cited in the application the whole document	
A	M P SCHLUNEGGER ET AL: "An unusual feature revealed by the crystal structure at 2.2 Å resolution of human transforming growth factor-beta 2" NATURE,GB,MACMILLAN JOURNALS LTD. LONDON, vol. 358, no. 358, page 430-434-434 XP002110204 ISSN: 0028-0836 cited in the application the whole document	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 00/07600

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9417099 A	04-08-1994	US 5420243 A	30-05-1995
		AU 697916 B	22-10-1998
		AU 4097797 A	08-01-1998
		AU 685084 B	15-01-1998
		AU 6093294 A	15-08-1994
		CA 2153789 A	04-08-1994
		DE 69424576 D	21-06-2000
		EP 0791011 A	27-08-1997
		ES 2146253 T	01-08-2000
		JP 8510443 T	05-11-1996
		US 5658883 A	19-08-1997
WO 9504819 A	16-02-1995	DE 4420157 A	23-02-1995
		AT 189475 T	15-02-2000
		AU 688362 B	12-03-1998
		AU 7498694 A	28-02-1995
		CA 2169171 A	16-02-1995
		CN 1129013 A	14-08-1996
		CZ 9600357 A	17-07-1996
		DE 59409126 D	09-03-2000
		DK 713529 T	19-06-2000
		EP 0713529 A	29-05-1996
		ES 2142953 T	01-05-2000
		GR 3032628 T	31-05-2000
		HU 74271 A	28-11-1996
		JP 9501053 T	04-02-1997
		NZ 271376 A	24-04-1997
		PT 713529 T	30-06-2000
		SI 713529 T	30-04-2000
		US 5994094 A	30-11-1999
		ZA 9405992 A	14-03-1995
WO 9601316 A	18-01-1996	DE 19511243 A	04-01-1996
		AU 2979895 A	25-01-1996
		CN 1151758 A	11-06-1997
		DE 19580745 D	11-03-1999
		JP 10502527 T	10-03-1998
		US 6120760 A	19-09-2000
		US 5807713 A	15-09-1998
WO 9516035 A	15-06-1995	ZA 9505444 A	14-02-1996
		US 6027919 A	22-02-2000
		AU 689184 B	26-03-1998
		AU 1301395 A	27-06-1995
		CA 2176942 A	15-06-1995
		EP 0733109 A	25-09-1996
		FI 962350 A	16-07-1996
		JP 9506261 T	24-06-1997
		KR 249923 B	15-03-2000
		NO 962304 A	04-06-1996
		US 5658882 A	19-08-1997
EP 0915168 A	12-05-1999	JP 9295945 A	18-11-1997
		AU 2408497 A	19-11-1997
		NO 985041 A	29-10-1998
		CZ 9803449 A	17-03-1999
		WO 9741250 A	06-11-1997
		PL 329610 A	29-03-1999

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP 00/07600

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9703188 A	30-01-1997	DE 19525416 A	16-01-1997
		AU 6615196 A	10-02-1997
		EP 0837938 A	29-04-1998
		JP 11509097 T	17-08-1999

**THIS PAGE BLANK (USPTO)**

---